

Quasi-Realism and the Moral Problem

Master thesis in Philosophy

Andreas Brekke Carlsson

Department of Philosophy, Classics, History of Arts and Ideas

UNIVERSITET I OSLO

November 2007

Acknowledgements

First I want to thank my supervisor Professor Christel Fricke. She has provided valuable feedback and support for which I am very grateful. I also want to thank my friends Mikkel Storm Glomstein, Torfinn Huvenes, Knut Olav Skarsaune, Hallvard Stette and Tore Øgaard who all took time to read and comment on the text. Furthermore I am grateful to my family for their firm support of my philosophical endeavours. Last but not least my thanks go to Katarzyna Kuś who not only read and commented the text but also has supported me in numerous other ways.

Table of Contents

1. Introduction.....	5
1.1 The Moral Problem.....	5
1.2 The Expressivist Solution to the Moral Problem.....	7
1.3 Problems with the Expressivist Solution.....	9
1.4 The Quasi-Realist Solution.....	11
1.5 Aim and Structure of the Thesis	12
2. Motivational Internalism.....	15
2.1 Motivational Internalism.....	15
2.2 Do Psychopaths Make Moral Judgments?.....	17
2.3 Does Satan make Moral Judgments?.....	19
2.4 Other Counterexamples.....	21
2.5 Smith's Argument for Internalism.....	22
2.6 Expressivism and Weakness of Will.....	24
3. The Humean Theory of Motivation.....	29
3.1 The Humean Theory of Motivation.....	29
3.2 An Anti-Humean Argument from Moral Phenomenology.....	31
3.3 A Dispositional Conception of Desire and the Teleological Argument.....	32
3.4 Can Directions of Fit Differentiate Beliefs and Desires?.....	34
3.5 Motivated and Unmotivated Desires.....	38
3.6 Besires.....	39
3.7 Summary and Preview.....	42
4. Moral Attitudes and the Demarcation Challenge.....	44
4.1 Attitudes and Normativity.....	44
4.2 Higher Order Attitudes.....	45
4.3 Stable Attitudes.....	48
4.4 Prescriptive Attitudes.....	51
4.5 Moral Authority and Justification.....	53

5. Mind-dependence and the Internal Reading.....	55
5.1 Error and Mind-Dependence.....	55
5.2 The Internal Reading.....	59
5.3 The Internal Reading and Relativism.....	61
5.4 The Internal Reading and Generality.....	65
 6. Quasi Realism and Moral Fallibility.....	 68
6.1 The Errors of Others and the Errors of my own.....	68
6.2 Passing the Buck.....	69
6.3 Blackburn on Moral Fallibility.....	71
6.4 Egan's Argument.....	74
6.5 Biting the Bullet.....	78
 7. Summary and Conclusions.....	 80
 Bibliography.....	 84

1. Introduction

1.1 The Moral Problem

1.2 The Expressivist Solution to the Moral Problem

1.3 Problems with the Expressivist Solution

1.4 The Quasi-Realist Solution

1.5 Aim and Structure of the Thesis

1.1 The Moral Problem

In *The Moral Problem*, Michael Smith describes two characteristic features of moral judgments. First, they aim at objectivity. When we ponder a moral question, we normally think there is some kind of correct answer to be found. This answer does not depend upon our own reasoning or feelings. It will be the right thing to do whether or not we think it is. The fact that somebody claims that something is right does not make it right. This feature of morality is manifested in the way we deal with the possibility of moral error. Not only do we accuse others of making moral errors, we also wonder whether our own judgments are correct or not. One natural way of formulating this idea about the possibility of correct moral judgments is to say that there are moral facts. Just as there are facts about what I do, there are facts about what I ought to do. Thus, moral judgments attempt to capture independently existing moral facts, and when they do, these judgments are true. In this way moral judgments function as ordinary beliefs. We say things like: “Capital punishment should be abandoned.” This judgment seems to be capable of being either true or false, and it might possibly constitute moral knowledge.

The second feature is that moral judgments are essentially practical. When we make moral judgments we do not merely aim to capture some truth about how the world is, we also take these judgments to guide our conduct. Imagine that I am at a restaurant with a friend. While we enjoy our soup we embark on a long discussion on vegetarianism and by the time we are ready to order the main dish, my friend has convinced me that it is wrong to eat meat. If I then go on to order a steak there will obviously be something wrong with me. I will either be taken to be insincere in my judgment, weak willed, or suffering from some strange meat-eating compulsion. So changes in our moral views reliably track changes in our motivation. Of course, we do not always act on our moral judgments, but if we do

not have any motivation at all to act on our moral judgments, people will rightfully question our sincerity.

The two features I have considered can be summarized as follows. Moral judgments aim at representing some facts about the world. But they also function to prompt actions and effect changes in the world. But if we accept a very plausible theory of human psychology, the one developed by David Hume (1888) these two features of moral judgments seem incompatible. According to Hume, we can divide our mental states into two categories. On the one hand there are beliefs; states of mind which aim to represent the world. On the other, there are the appetitive states of mind; needs, desires, attitudes and emotions. These states of mind aim to prompt action and effect changes in the world. So beliefs and desires are characterized by different functional roles. A belief alone is inert; it needs to be supplemented by a desire in order to motivate. A desire on the other hand cannot represent aspects of the world; it is neither true, nor false. These two states are, according to Hume, distinct existences. It is always possible to have a belief without having the appropriately related desire and *vice versa*.

The problem with moral judgments is that they share both the features of beliefs and those of desires. Like beliefs they aim at objectivity and truth, like desires they are capable of motivating us. But according to the Humean psychology, moral judgments must either be the expressions of beliefs, or the expressions of some desire-like state. If we opt for the first alternative we seem incapable of explaining the practicality of moral judgments. If we opt for the second alternative we seem to be forced to give up the objectivity of moral judgments. This dilemma is what Michael Smith calls the moral problem.

There are three main ways of tackling this dilemma. The first solution is to deny that there is any necessary connection between moral judgments and motivation. In this case moral judgments could still be beliefs; whether we are motivated by these beliefs is however a contingent matter. The second solution is to maintain that there is a necessary connection, but deny the Humean claim that beliefs cannot motivate. In this thesis, however, I want to focus on the third approach. This is to hold on to both the claim that moral judgments are necessarily motivating, and the Humean psychology, but to deny that moral judgments aim at describing or representing the world. This is the expressivist solution to the moral problem.

1.2 The Expressivist Solution to the Moral Problem

According to expressivists, moral judgments do not express beliefs about how the world is.¹ What we do when we make a moral judgment is rather to express an attitude, or an emotion towards an action, or situation.² In *Language, Truth and Logic*, A. J. Ayer gives a famous description of this position:

Thus if I say to someone, “You acted wrongly in stealing that money”, I am not stating anything more than if I had simply said, “You stole that money.” In adding that this action is wrong I am simply evincing my disapproval of it. It is as if I had said, “You stole that money” in a peculiar tone of horror, or written it with the addition of some special exclamation marks. (Ayer: 1946: 107)

Whereas beliefs have a representative function, emotions, desires or attitudes do not. In expressing a belief I am trying to say something about the way the world is. This is not the case with expressing an emotion. When I say “Hooray for Manchester United,” I am not saying something that can be true or false about the world. Instead I am expressing my feelings towards a football club. According to Ayer, the same structure holds for moral utterances. When I say that stealing is wrong, I am not expressing the belief that there are facts about stealing that makes it wrong; I am just expressing my disapproval towards stealing. An expression of emotion (or an attitude, or a disposition) cannot be either true or false. Moral utterances, according to Ayer, are not truth-apt.³ Understanding moral judgments in terms of descriptions and beliefs, expressivists claim, will imply a too narrow

¹ I will use the term “expressivism” when referring to the metaethical tradition from Ayer (1946), Stevenson (1944) and Hare (1952) to Blackburn (1984; 1993; 1998) and Gibbard 1990; 2003). “Non-cognitivism” is also used in the literature, but as Blackburn wants to distance himself from this tag (Blackburn: 1996), I will stick to “expressivism.”

² Except from the claim that it is not a cognitive state, expressivists differ concerning which psychological state is being expressed by a moral sentence. Ayer opted for emotions; hence emotivism; Hare spoke about prescriptions; hence his prescriptivism. Blackburn does not want to commit himself to any of these and speaks loosely about attitudes. I will return to this question in chapter 4, but for now I will adopt Blackburn’s use of “attitudes” when I am speaking of expressivism in general.

³ Ayer and later expressivists are careful to distinguish emotivism from subjectivism. Subjectivism holds that when I am saying that stealing is wrong I am claiming that I have the feeling that stealing is wrong. Whereas emotivism claims that moral sentences consist in *expressing* my feelings towards an action or a situation, subjectivism holds that I am *reporting* my feelings. This is an important difference. If the function of the sentence: “Stealing is wrong” is to report that I disapprove of stealing, the sentence would have truth conditions. It would be true if I actually had a disapproving feeling towards stealing and false if I did not have it. But on this picture it will be impossible to account for moral disagreement. If Smith says “stealing is wrong” and Jones says “stealing is not wrong”, they are not, according to subjectivism, contradicting each other. They are merely stating their beliefs about their respective emotions towards stealing.

and too theoretical conception of morality. Ethics is essentially practical. The function of morality is to guide conduct, coordinate actions and to lay pressure on others. In this way they accommodate the motivational aspect of moral judgments. It is uncontroversial that attitudes, desires and emotions are able to motivate. Before turning my attention to the question of moral objectivity I will briefly mention some of the strengths of this position.

Because expressivists understand moral judgments as the expressions of attitudes, they deny that there are any moral facts, or moral truths. Expressivism is therefore a variant of moral anti-realism. The denial of moral facts fits well with a naturalistic description of the world. The term “moral facts” can be given either a naturalistic or a non-naturalistic interpretation. On the non-naturalistic interpretation moral realism faces a difficult problem consisting of two components, one metaphysical and one epistemological. This problem is succinctly formulated by John Mackie. First, non-natural moral facts seem metaphysically *queer*. “If there were objective values then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe” (Mackie 1977: 38). The second part of the problem has to do with our knowledge of such facts. Since the moral facts are not to be understood as natural facts, it is puzzling how we could have an epistemic access to them. “Correspondingly, if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else” (Mackie 1977: 38).

Expressivism escapes these problems. On the expressivist account, there is nothing more to moral judgments than human beings projecting their emotions, attitudes or dispositions onto the “disenchanted, non-ethical order we inhabit” (Blackburn 1998: 49). The states of minds that are projected are themselves ordinary psychological states and therefore proper objects for scientific study. Expressivism is thus a strictly naturalistic theory which claims to give a comprehensive explanation of ethics without evoking any non-naturalistic facts or properties. This denial of moral facts does not only fit well with a scientific description of the world; it also gives the theory a better metaphysical and epistemological economy. As Simon Blackburn puts it: “The projective theory intends to ask no more from the world than we know is there [...] It asks no more than this: a natural world, and patterns of reactions to it” (Blackburn 1984: 182).

But it is not necessarily the case that moral facts must be understood as mystical, platonic entities. It is possible to be a moral realist and yet maintain a naturalistic image of the world. One⁴ way of doing this is to embrace definitional naturalism, which claims that the meaning of moral terms such as “good”, or “right” is equivalent to some naturalistic description such as “conducive to happiness”. According to definitional naturalism “X is good” means the same as “X has the natural property F”. This form of naturalism is traditionally countered by G. E. Moore’s “Open Question Argument” (Moore: 1903: 64-69). Take the claim that “X is good” means the same as “X is conducive to happiness.” According to definitional naturalism this is an analytical truth; it would be self-contradictory to claim that X is good, but that X is not conducive to happiness. But it is not self-contradictory for a competent English-speaker to claim this. So, it cannot be an analytical truth that “X is good” means “X is conducive to happiness”. The same Open Question Argument can be applied to whatever naturalistic property “good” is claimed to be synonymous with. The soundness of this argument has been the subject of much debate. My point here is only that it does not apply to expressivism, or any form for anti-realism. If there are no moral facts, the question of whether or not moral facts are reducible to natural facts does not arise. According to the expressivist moral predicates cannot be defined in non-moral, descriptive terms because moral terms do not refer to properties; all they do is to express the attitudes of the speaker. If the “Open Question Argument” is sound, it need not be taken to show that moral terms describe some *sui generis*, non-natural facts, as Moore argued. Rather, the expressivist would claim, moral terms do not describe facts at all.

1.3 Problems with the Expressivist Solution

In the section above I gave a broad overview of some of the reasons why the expressivist solution to the moral problem appears attractive. Not only is the theory capable of giving a full account of the practical, action-guiding character of morality, it also promises an explanation of ethics that does not face the

⁴ Definitional naturalism is of course not the only naturalistic version of moral realism. But I do not have the space to go into other theories here.

metaphysical and epistemological difficulties which realist theories must overcome.

But given its denial of moral objectivity, expressivism gives an extremely revisionist account of morality. Ordinary moral thought and language exhibit a lot of features which seem at odds with expressivism. Moral sentences seem to be fact-stating. We speak of wrongness and rightness of actions, or of good and evil deeds, as if we were describing facts and properties of the world. We also speak of moral truths, of moral beliefs and moral knowledge. This is not just a linguistic matter; it exemplifies common and ordinary ways of thinking about morality. Our ordinary conception of morality includes claims like: i) Some moral judgments are true; ii) That someone thinks that an action is right does not make it right; it is possible to make moral errors; iii) Not every moral judgment is equally good, or equally justifiable. The phenomenology of morality also point in the direction of objectivity. Moral obligations are commonly perceived as stemming from an external source and not from our own feelings or attitudes; morality has “objective feel.” (Blackburn 1993: 153, Timmons 1999: ch. 3) This authority of moral claims disappears on the expressivist interpretation. To understand moral judgments in terms of desires and attitudes seems to degrade morality into a mere matter of taste. Some people have a positive attitude towards capital punishment, some have a negative, but just as with matters of taste the question of who is *really right* does not admit of an answer. In moral questions this lack of objectivity is deeply disturbing.

Ayer accepted this revisionist consequence of expressivism. Moral judgments, he claimed, “have no objective validity whatsoever.” (Ayer 1946:108) Early expressivism presented a picture of morality that fitted nicely with a naturalistic view of the world. To do this however, it had to deny so many commonly held assumptions that it became a kind of error theory⁵. But the claim

⁵ This kind of error theory differs from John Mackie’s (1977). Mackie’s theory is based on a conceptual and an ontological claim. The conceptual claim is that “ordinary moral judgements include a claim to objectivity, an assumption that there are objective values” (Mackie 1977: 35). The ontological claim is moral anti-realism: the denial of the existence of objective values. There is thus a mismatch between what we are claiming to be true when we are making moral judgments and what actually is true; hence, moral judgements are uniformly false. The difference between error theory and expressivism is often said to be that while expressivists accept Mackie’s ontological claim, they deny his conceptual claim: What we are doing when we are making moral judgements is not to assume some objective values, but to express attitudes. Moral terms do not refer to moral properties or fact because they are not descriptive at all. Nonetheless early expressivism collapses into a different kind of error theory because it denies a large amount of commonly held assumptions of moral objectivity.

that there is something globally wrong with ethics is not a very attractive position, given ethics' fundamental role in our life. So, the expressivist solution to the moral problem is very unsatisfying. Instead of a solution to the problem, it appears as a way of escaping it; the dilemma remains.

1.4 The Quasi-Realist Solution to the Moral Problem

One way out of this dilemma is to argue that expressivism does not imply a revisionary theory of morality. With the introduction of the position called quasi-realism, Simon Blackburn presents a version of expressivism, which, while holding on to some of the main tenets of Ayer's emotivism, does not give up on our commonsensical ideas of moral objectivity.⁶ What distinguishes quasi-realism from earlier versions of expressivism is that it explicitly aspires to be non-revisionary.⁷ Quasi-realism is the project of explaining and justifying ordinary moral thought and language on a slender and austere metaphysical basis. One part of this project is semantic in nature. Blackburn aims to show how the claim that moral judgments express attitudes can be married with the fact that moral judgements are expressed in sentences which share the syntactical features of expressions of beliefs.⁸ I will leave this part of the quasi-realist project aside in order to focus on issues related to moral psychology and moral objectivity. The other part of the quasi-realist project aims to show that moral anti-realism, the thesis that there are no moral facts or properties, is compatible with our commonsensical understanding of morality. The theoretical advantage of this theory is obvious. Whereas traditional expressivism implies that we are mistaken about our moral judgments, quasi-realism, on the other hand, "is well seen as the

⁶ Quasi-realism is therefore a subspecies of expressivism. I will speak of "expressivism" when I discuss arguments which relates generally to the expressivist claim that moral judgments express attitudes (mostly in chapters 2 and 3). I will speak of "quasi-realism" when I discuss arguments that relates specifically to Blackburn's non-revisionary brand of expressivism (mostly in chapters 4, 5 and 6).

⁷ For similar non-revisionary versions of expressivism, see Gibbard (1990); (2003) and Timmons (1999). I choose to focus on Blackburn's quasi-realism because of these authors Blackburn is most explicit in his concern for saving the objectivity of moral discourse.

⁸ We normally express our moral judgements in indicative sentences and not as exclamations or commands. Moral sentences can be embedded in the ascription of propositional attitudes. Moreover, moral judgements may appear in logical complex contexts, for example as antecedents of conditionals and may appear as premises in valid arguments. How an expressivist can explain the semantics of moral utterances when they appear in non-assertive contexts, for example as the antecedent of a conditional is one of the most debated problems in the literature. The problem was first presented in Geach (1960; 1965). Blackburn's two main attempts to solve it is found in Blackburn (1984: 189-196) and (1993: 182-197). It remains highly controversial whether any expressivist attempts to solve this problem have proven successful. For criticism, see for example Hale (1986) and (1993).

attempt to save expressivism from error theory. It attempts to show that ordinary moral thought is not affected root and branch with philosophical myth” (Blackburn 2006a: 154).

Quasi-realism consists of two components, an explanatory and a justificatory story. The explanatory story seeks to place morality well within the domain of naturalism by giving an interpretation of moral judgments as the expressions of attitudes. In this respect it follows the various forms of expressivism developed by Ayer (1946), Stevenson (1944) and Hare (1952). But quasi-realism is also offering a justificatory story, showing how the objective features of morality can be maintained even given the expressivist explanatory story. This story is the reason why quasi-realism is news, and why it promises to justify our ordinary notions of morality at far lower metaphysical costs than moral realism. Blackburn’s strategy in achieving this goal is to “domesticate” the features associated with moral objectivity, such as moral mind-independence, moral error and moral anti-relativism.

“These concerns and claims look metaphysical, and indeed many theorists are content to define their metaethical theory in terms of them. But looked at another way they are merely part of good ethical thought: someone incapable of them would lack a becoming modesty, rather than a metaphysical insight” (Blackburn 1993: 4).

The interesting question which quasi-realism brings to the surface is therefore this: How much metaphysics is needed to sustain our commonsensical thoughts about morality? If the quasi-realist project is successful the answer would be: very little.

1.5 Aim and Structure of the Thesis

The overall aim of this thesis is to give a discussion of quasi-realism as an answer to the moral problem. Following Smith (1994a), we can formulate this problem in terms of three independently plausible claims. Taken together, however, they appear to be inconsistent.

1. Moral judgments a) aim at objectivity, so b) moral judgments are beliefs.

2. Moral judgments have a necessary connection to motivation.

3. In order to be motivated we need both a relevant desire and a means-end belief. A belief alone is not enough to motivate, and beliefs and desires are distinct existences; so no belief has a necessary connection to motivation.⁹

Given the Humean theory of motivation, the two characteristic features of moral judgments; their aim at objectivity and their practicality, seem incompatible. Blackburn attempts to solve this dilemma by arguing that although moral judgments are expressions of attitudes, the quasi-realist can still justify the objective pretensions of moral discourse. He therefore embraces 1a) but denies 1b). This is not an easy task. Moreover, it might not be necessary. Both 2) and 3) have been contested. If it can be shown that either i) the connection between moral judgments and motivation is contingent, or ii) that beliefs alone can motivate, the rationale for the quasi-realist solution disappears. In order to give a proper discussion of the quasi-realist answer to the moral problem, two matters must be investigated. First, whether 2) and 3) provides us with a good reason to accept the quasi-realist solution to the moral problem. Second, whether quasi-realism achieves what it promises: to save the objective features of morality. That is, does he succeed in rejecting 1b without losing 1a?

The two first chapters will discuss motivational internalism and Humean theory of motivation. Taken together, these two positions seem to yield a knock-down argument for expressivism. I will argue that although there are good reasons for accepting internalism, the most plausible formulation of it is in fact incompatible with quasi-realism, or any kind of expressivism. Moreover, the Humean theory of motivation faces some grave problems on its own. So the apparent knock-down argument is not available to Blackburn. Still, if it can account for the objective features of morality, quasi realism will be very attractive theory because it promises to account for moral objectivity without appealing to any heavy duty notions of moral facts or properties. In the last three chapters I discuss Blackburn's attempts to show that quasi-realism does not imply any kind of moral revisionism. In chapter 4, I will discuss his attempt to demarcate moral attitudes from non-moral attitudes. Although moral judgments motivate, not every

⁹ This is a modified version of Smith's formulation of the moral problem (Smith 1994a: 12)

motivational state can count as a moral attitude. From the discussion in this chapter it will be clear that Blackburn needs some account of how our moral attitudes can be justified. The two last chapters will discuss some obstacles for this project. First, Blackburn's account of moral judgments has been taken to imply that morality is mind-dependent. In chapter 5 I aim to defend Blackburn against this accusation. Chapter 6 will deal with the problem of moral fallibility. Although Blackburn can account for how others may be wrong about ethical matters, it is far more difficult, on the quasi-realist framework, to give an account of the possibility that we ourselves may be in error. This is, in my opinion, where the quasi-realist project breaks down. I will argue that although Blackburn's quasi-realism improves on its expressivist predecessors by giving a very subtle account of how many features of moral objectivity can be reconciled with moral anti-realism, the project ultimately fails because it is unable to give a plausible account of moral fallibility.

2. Motivational Internalism

- 2.1 Motivational Internalism
- 2.2 Do Psychopaths Make Moral Judgments?
- 2.3 Does Satan make Moral Judgments?
- 2.4 Other Counterexamples
- 2.5 Smith's Argument for Internalism
- 2.6 Expressivism and Weakness of Will

2.1 Motivational Internalism

Moral judgments play a fundamental practical role in our life. One way of bringing out this practical function of moral judgments is by noting the close connection between moral judgments and actions. Moral judgments have a pull on us. They seem to be intrinsically motivating. There is something wrong with a person who claims that one ought to give to the homeless, but all the same strides past them every day on his way to work without opening her wallet; we will suspect such a person of being insincere. Although it is intuitively plausible that there is some kind of connection between moral judgments and actions, it is not entirely easy to pin down just what this connection consist in. According to *motivational internalism*, there is a necessary connection between moral judgments and motivation. According to *motivational externalism*, on the other hand, this connection is merely contingent. As noted in the introduction, expressivists are internalists. But internalism comes in two different versions, as strong motivational internalism (SMI), and as weak motivational internalism (WMI)

SMI: If an agent judges that it is right for her to ϕ in circumstances C, then she is motivated to ϕ in C¹⁰

WMI: If an agent judges it is right for her to ϕ in circumstances C, then she is motivated to ϕ in C, except in cases where the agent is suffering from motivational disorders that affect her more generally.¹¹

¹⁰ I owe this formulation to Smith (1994a: 61).

¹¹ This formulation is a modified version of Svavarsdottir's (1999) preferred formulation of internalism.

Neither strong nor weak internalists commit themselves to the implausible claim that people always act in accordance with what they judge to be right or good. The claim is rather that if an agent judges that she ought to ϕ , then she will have some kind of motivation to ϕ . There might of course be other concerns which override this motivation. Although I am motivated to relieve the suffering of the people of Africa, I am also motivated to buy a new stereo, and this motivation might well prove to be the stronger of the two. But this does nothing to undermine motivational internalism as stated above.

The difference between SMI and WMI is that the latter states a *defeasible* connection between moral judgments and motivation. We can imagine, and indeed experience, situations where we, because of weakness of will, or depression, judge something to be morally good, but all the same fail to be motivated. This is the reason why WMI is a far more popular position than SMI. While it is commonly assumed that internalism needs some sort of clause to allow for weakness of the will, it is controversial how this clause is to be formulated. Michael Smith (1994a) defines WMI in the following way: “If an agent judges that it is right for her to ϕ in circumstances C, then either she is motivated to ϕ in C, or she is practically irrational.” The problem with this formulation is that Smith is not very precise when it comes to specifying what practical irrationality amounts to. He writes that it should cover cases of “weakness of the will and the like” (Smith 1994a:120). But as Alexander Miller has pointed out (Miller 2003: 221), there is a danger that “and the like” is interpreted as any condition that would break the connection between moral judgment and moral motivation. This would make Smith’s formulation trivial. The formulation would not allow for the possibility of counterexamples. In order to overcome this difficulty I have adopted a suggestion from Sigrun Svavarsdottir (1999) and define WMI in terms of motivational disorders that affect the agent more generally. This formulation escapes the problem of triviality because the “motivational disorders” are to affect the agent’s motivation in a general way, and not merely her moral motivation. It allows for the possibility of counterexamples to internalism. If a person makes a moral judgment and yet fails to be motivated, this will count as a counterexample to WMI, as long as the lack of motivation is not a product of a general motivational disorder that manifests itself in other parts of the agent’s life.

In the next four sections of this chapter I will argue for the plausibility of WMI. (Since this is the most common version of internalism I will just refer to WMI as “internalism” until the last section of this chapter). This does however not amount to an argument in favour of expressivism. For, as I will argue in section 2.6, any version of expressivism will face grave difficulties in accounting for how WMI can be true if moral judgments are expressions of non-cognitive attitudes. The only version of internalism open for the expressivist is strong motivational internalism.

2.2 Do psychopaths make moral judgments?

One of the main arguments against any kind of motivational internalism is the possibility of an amoralist, a person who makes a moral judgment and yet remains unmoved, although not suffering from motivational disorders (Brink 1989). It is important to note that the judgment must be a genuine moral judgment for the case to count as a counterexample to internalism. The expressivist admits that we sometimes use evaluative terms in the “inverted comma sense”; not as expressions of our own attitudes but rather to point out that a certain action meets the standards we normally use for an evaluative term (Hare 1952: 124; Blackburn 1998:60). We might for instance claim that it is good to give to charity, but thereby only expressing that giving to charity is the kind of thing people usually call “good.” This is a way of expressing moral conventions. But this is not a problem for the internalist. She can simply answer that these cases are not examples of genuine moral judgments. So the question is whether someone can sincerely claim that an action is morally good without being motivated.

Psychopaths may provide an example. The mass murder Robert Harris for instance, claimed that he knew that his actions were morally wrong. He just did not feel motivated to act in accordance with what he judged to be morally good (Watson 1987). In this case, it seems strained to account for the lack of motivation on Harris’ behalf as weakness of will or depression. The internalist answer to examples such as Harris is to deny that he is actually making a moral claim. If he is not motivated by the fact that he judges murder to be wrong, he is not really making a moral judgment. Rather he is using moral terms in the inverted comma

sense; he is merely acknowledging that his actions belong to a category which is conventionally labelled “wrong”.

There is some evidence to back up this claim. In *Sentimental Rules*, Shaun Nichols discusses recent psychological studies on psychopaths (Nichols 2004: ch.1 and 3). One characteristic feature of this group is that they fail to distinguish between morally wrong and “conventionally wrong.” This difference is not always straightforward. All the same, studies¹² show that children from a young age are able to distinguish paradigmatic cases of “morally wrong” such as hitting another person, pulling hair, pushing or stealing, from paradigmatic cases of conventional wrongness such as violation of school rules, or violation of etiquette, for instance talking out of turn or drinking soup out of a bowl. Children tend to claim that moral transgressions are more serious, and wrong not only in one country or school but also in other countries and other schools. Conventional transgressions on the other hand are taken to be less serious and more contingent upon authority. Hence, if the teacher said it was allowed to chew gum in class, it would be permissible to do so. But if the teacher said it was allowed to hit other children this would not be permissible.

In a series of studies Blair (1995) presented imprisoned psychopaths and a control group of imprisoned non-psychopaths with similar tasks. Whereas the control group distinguished between moral and conventional transgressions, the psychopaths did not. The psychopaths evaluated the conventional transgressions to be as serious and non-permissible as the moral ones. Moreover, the two groups diverged in their explanation of why the actions in the examples were wrong. The non-psychopath tended to explain the wrongness of the moral transgressions by appealing to the victim’s welfare. The psychopath however, gave conventional-type explanations for the moral as well as for the conventional transgressions. As Nichols (2003: 76-77) points out, these studies give us some reason to claim that psychopaths are not really making moral judgments, but are rather using moral terms in an inverted commas sense. If this is correct, psychopaths do not provide us with a counterexample to motivational internalism.

To claim that psychopaths do not make genuine moral judgments might seem to beg the question against the externalist. Motivational internalism claims

¹² See Blair (1995).

that moral judgments are intrinsically motivating. Psychopaths are not motivated. Hence, they are not making any moral judgment. Construed this way, the argument that psychopaths do not make genuine moral judgments fails. But this would be the case only if motivation was the only thing the psychopaths lacked. Blair's studies, however, provide us with independent arguments that point to the conclusion that psychopaths are not making genuine moral judgments. They do not only lack the motivation to act in accordance with what they claim to be good, but also the ability to treat moral transgressions as a distinct class. Arguably, this ability is a necessary condition for passing moral judgments.

2.3 Does Satan make moral judgments?

But what about cases where people judge something to be good but still does the opposite? We do seem capable of acting intentionally against our own values. Such cases cannot be described as psychopathic behaviour. First, because cases like these are far too common. Second, because "values" is here to be understood as genuine values rather than in the inverted comma sense. Indeed, one explanation would be to deny that people can intentionally act against their own values. In that case they merely use "good" in an insincere manner. But this claim is based on a simplistic psychology. Take for example the perhaps most complex of Dostoyevsky's characters, Stavrogin, in *The Demons*. After he seduces a young girl, who subsequently kills herself, he commits a series of evil acts, not because he "really" takes them to be good, but exactly because they are evil. This is an extreme, but nonetheless possible psychology.

A natural reply to cases like these would be to claim that our desires often are stronger than our moral convictions. The spirit is willing but the flesh is weak. Although we know what is good, we often succumb to all kinds of temptations. This commonsensical explanation does account for many, and perhaps most, of the cases where we act against our values. But it does not explain the cases where people act against their values *because* it is against their values, or do something evil *because* they take it to be evil. In *Ruling Passions*, Blackburn (1998) uses Milton's Satan as an example. In *Paradise Lost* Satan famously claims: "Evil, be thou my Good." Satan's judgment that something is evil is therefore what attracts

him to do it. Do examples like this give the externalists the upper hand? Blackburn argues that they do not:

“[Externalists] can certainly point to possible psychologies about which the right thing to say is that the agent knows what it is good or right to do, and then deliberately and knowingly does something else [...] But internalists win the war for all that because these cases are necessarily parasitic, and what they are parasitic upon is a background connection between ethics and motivation. They are cases where things are out of joint, but the fact of a joint being out presupposes a normal or typical state in which they are not.” (Blackburn 1998: 61)¹³

Blackburn’s point is that cases like Satan or Stavrogin can only be understood as *conflicts* within a person. Both Stavrogin and Satan are despairing figures. Psychologies like these do not show that moral judgments are inert, or that the relationship between moral judgments and motivation is merely contingent. To do evil because it is evil presupposes some prior motivation to do the good. The evil deeds of Satan and Stavrogin should be understood as *reactions*, either against the agent himself or some of his internalized concerns for what is good. This is clear in the case of Stavrogin. Because he cannot forgive himself, he commits evil acts in order to punish himself. But the reason why these acts are intelligible as self-inflicted punishment and not merely as sadism is some motivation to do what is good. Satan, on the hand, does not punish himself, but he is “racked with all the “foul distempers” going” (Blackburn: *ibid*).

To judge something as evil (and not merely to use “evil” in the inverted comma sense) presupposes knowledge of what is good. And a characteristic feature of people who do evil deeds because they are evil is a conflict within this person. But in order to make sense of this conflict, the knowledge of the good, which the evil deed presupposes, has to be conceived as an active, and not an inert state of mind. It is in this sense that Stavrogin’s and Satan’s abnormal

¹³ Dreier (1990: 11) makes a seemingly similar point: “I am inclined to say that the failure of internalism is parasitic on a background of its success. The internalist will say that the exceptions to his principle seem to reside around the edges of logical space, whereas the important cases are safely in its core” But the similarity is merely apparent. Dreier allows for cases where people are not motivated by their moral judgment, but argues that these cases are not normal. Our *understanding* of cases like these is thus parasitic upon a notion of normality. This is problematic because, as Dreier himself admits, he needs an independent notion of normality, which he is not able to provide. Blackburn’s point is not that our understanding of Satan is parasitic upon some notion of normality, but rather that Satan’s state of mind itself is parasitic upon a normal state of affairs. Whereas Dreier invokes “normality” to make sense of our internalist intuitions, Blackburn uses a normal state of affairs to make sense of Satan’s psychology.

psychologies are parasitic upon a normal state, where knowledge of the good motivates.

2.4 Other counterexamples

It is perhaps because of arguments like Blackburn's that psychologies like Satan's or Stavrogin's are not typical of the proposed counterexamples offered by externalists. Rather than focusing on people who willingly and knowingly do what they consider to be bad, these counterexamples typically describe persons who believe that something is good but nevertheless fail to be motivated by this belief. In *Moral Realism* (2002 148-155), Russ Schafer-Landau provides us with a case of this sort. He asks us to imagine a soldier who in two weeks time will be called to the front. Dwelling on the horrors of war, he loses all his motivation to fight. But he still considers it to be his duty to go to war. He has no motivation whatsoever to do what he considers to be his duty. Given that this is how the soldier conceives of his situation the internalist must ascribe to him some mistake in order to show that this is not a genuine counterexample. As Schafer-Landau points out, there are three kinds of mistakes that the agent might commit.

“(i) The agent may wrongly think that she is making a moral judgment – her statement may lack moral content. (ii) Her statement may have moral content, but she doesn't really endorse it, though she thinks that she does. (iii) She may commit neither of these mistakes, but instead fail to notice that she is, after all, motivated to comply with her sincere moral judgment.” (Schafer-Landau 2003: 152)

The two last mistakes would imply that the soldier is mistaken about his own psychology. According to ii) he mistakenly thinks that he endorses the duty to fight. According to iii) he fails to recognize his motivation to fight. Schafer-Landau admits that agents do make mistakes of this kind. But his point is that internalists cannot rely on such mistakes in order to discredit the counterexamples. We do not know whether the soldier makes these mistakes or not, so to claim that he does cannot be based on any knowledge of his psychology.¹⁴ Rather, to ascribe these mistakes to the soldier seems to follow from motivational internalism itself;

¹⁴ Sigrun Svavarsdottir (1999) makes a similar point.

the soldier must make mistakes like ii) or iii) because moral judgments do motivate. But this, of course, would be question-begging.

The best strategy for the internalist will therefore be to ascribe the soldier with the first kind of mistake. Although he thinks that he makes a moral judgment, he does not. The reason why this is a more promising solution for the internalist is that it does not presuppose that the soldier is wrong about his own psychology. He is right when he says that he does not have any motivation to fight. His mistake is rather about what a moral judgment is.

We can now see why counterexamples like this usually lead to a standstill. Internalists will claim that the case of the unmotivated soldier does not amount to genuine counterexamples because the soldier is not making a genuine moral judgment, whereas externalists will maintain that he does. Both sides can accuse the other of begging the question. The externalist will say that the internalist denies that the counterexamples are genuine because of the lack of motivation. But this would just to assume what is to be proven. Similarly, the internalist will claim that the externalist's counterexamples rest on the assumption that to use moral terms in a reliable way is enough to make genuine moral judgments. But this is exactly what the examples are meant to prove. As intuitions tend to get wobbly at this point in the dialectic, we need an independent argument to settle the question. This is what Michael Smith tries to do in *The Moral Problem*.

2.5 Smith's Argument for Internalism

Michael Smith points out that although motivational disorders do occur, there is a reliable connection between moral judgment and motivation. In a person who is not suffering from motivational disorders “a *change of motivation* follows reliably in the wake of a *change in moral judgment*” (Smith 1994a: 71). Suppose there is going to be a referendum on whether or not capital punishment is to be abolished. I am in favour of capital punishment and also motivated to vote against the abolishment of capital punishment. But as I discuss the matter with Jones, he convinces me that capital punishment is wrong. I am now convinced that I was fundamentally mistaken; capital punishment should be abolished. Given this change of my moral judgment, what happens to my motivation? If I am not suffering from motivational disorders, my motivations will change as well. This

reliable connection is something that externalists as well as internalists need to acknowledge. Smith's claim is that internalists can make sense of this fact, whereas externalists cannot.

When this reliable connection is explained on the internalist account "it follows directly from the content of moral judgment itself" (ibid: 72). For the expressivist, this is easy to explain. When I judge that it is right to vote for the abolishment of capital punishment, I express an attitude which itself is motivating.¹⁵

The externalist, however, cannot explain the reliable connection in terms of the content of the moral judgment itself because she holds that moral judgments are merely contingently motivating. But if the reliability of the connection is not to be explained in terms of the content of the moral judgment, it has to be explained by some facts about the person who is making the judgment. There has to be some disposition of the agent who is not suffering from motivational disorders that can track the connection. Smith calls this disposition for the content of moral motivation, as opposed to the content of moral judgment itself. But what can the content of the moral motivation be? Smith argues that the only possible motivational content that an externalist can evoke is a motivation to do the right thing, *whatever this might be*.

Whereas the internalist can explain the change in motivation that follows after a change in moral judgment as a direct consequence of the judgment, the externalist has to claim that the change in motivation is the product of two factors. Firstly, a motivation to do the right thing, whatever this might be, and secondly, my current judgment that something is right. On the externalist reading, my motivation to vote for abolishment of capital punishment is thus derived from my motivation to do the right thing, whatever that might be, together with the judgment that voting for abolishing capital punishment in fact is the right thing to do.

As Smith notes, this would give the externalist an explanation of the reliable connection. But there are two problems with this idea. First, it does not capture the phenomenology of motivation. When we consider what motivates us to some moral act, we normally do not find that the source of motivation was the

¹⁵ Smith's argument for internalism can be used by rationalists as well. In that case, "the belief that an act is right *produces* a corresponding motivation" (ibid).

desire to always do what is right. Rather, we find that we are motivated by our care for our friends, our love for our children and so on. Second, it conflicts with our commonsensical notions of a virtuous agent. The externalist explanation provides the virtuous agent with “one thought too many;” she becomes self-absorbed rather than virtuous. What motivates us to help our friends, care for our children and so on, *should not* be the desire to always do what is right, but rather our non-instrumental concern for our friends, our love for our children and so on. To postulate the desire to do what is right, whatever this might be, as the only source of motivation seems to alienate the agent from the proper concerns of morality. Hence, Smith claims that the “externalist explanation commits us to false views about the content of a good person’s motivations; it elevates a moral fetish into the one and only virtue” (ibid: 76).

2.6 Expressivism and Weakness of Will

I think that Smith’s argument makes weak motivational internalism an attractive position. But it is not at all clear that expressivists like Blackburn can use this form of internalism to show that moral judgments cannot be beliefs. The first problem is that moral beliefs might be motivating. If this is the case, the claim that moral judgments are necessarily connected to motivation does not provide the expressivist with an argument against cognitivism. This question is the topic of the next chapter. For now I want to address another worry concerning the relationship between internalism and expressivism. This will perhaps be puzzling since it seems obvious that expressivism can account for the motivating features of morality. But recall that internalism might be given two different interpretations, as strong motivational internalism (SMI), or as weak motivational internalism (WMI). In this chapter I have argued for the weak interpretation.

SMI: If an agent judges it right for her to ϕ in circumstances C, then she is motivated to ϕ in C.

WMI: If an agent judges it is right for her to ϕ in circumstances C, then she is motivated to ϕ in C, except in cases where the agent is suffering from motivational disorders that affect her more generally.

The problem with SMI is that it does not allow for cases of weakness of the will. But note that weakness of the will can be understood in two different ways. Normally we speak of weakness of will in cases where we make a judgment that I ought to do something but where I have desires that pull in a different direction. I make the judgment that I ought to stop smoking but my desire to have another cigarette is stronger. I do not think that I have a good reason to smoke, but I do it nevertheless. Cases like these, however, are perfectly compatible even with the strong version of internalism. From this version it only follows that I must have *some* motivation to do what I judge that I ought to do. Cases like the weak willed smoker only create difficulties for versions of internalism which hold that we always act on our moral judgments. But this would be a highly implausible variant of internalism.

The reason why WMI is far more popular than SMI therefore has to do with another form of weakness of the will. It is possible to become radically disaffected; to lose all motivation associated with moral judgments. These are cases where we hold something to be morally good, but have no motivation *whatsoever* to act on our judgment. In “Desiring the Bad: An Essay in Moral Psychology,” Michael Stocker (1979) claims that such cases indeed are commonplace:

“Through spiritual or physical tiredness, through weakness of the body, through illness, through general apathy, through despair, through inability to concentrate, through a feeling of uselessness or futility and so on, people may feel less and less motivated to seek what is good. One’s lessened desire need not signal, much less be the product of the fact that, or one’s belief that, there is less good to be obtained or produced, as in the case of a universal Weltschmerz. Indeed, a frequent added defect of being in such “depressions” is that one sees all the good to be won or saved and lacks the will, interest, desire or strength.” (Stocker 1979, 744)

Cases like these are the reason why Michael Smith among others opts for a *defeasible* connection between moral judgments and motivation. According to Smith we can make a sincere moral judgment “and yet seem to be totally indifferent to it; not motivated at all to do what we believe is good.” (Smith 2004: 142). This is not a form of externalism because this form of weakness of the will falls under the category of “motivational disorder that affects the agent more

generally.”¹⁶ Weak internalism is defined in a way that makes it invulnerable to Stocker’s counterexamples.

I will now argue that this weak version of internalism cannot be easily married with expressivism. According to the expressivist, to make a moral judgment is to express some non-cognitive state, a feeling, an attitude or some other motivating state. So when someone judges that it is right to ϕ in circumstances C, she will, at least there and then, have some motivation to ϕ . But what is the expressivist to say about Stocker’s counterexamples? The most natural reply would be to say that the severely depressed has *a belief to do what he thinks is right*, but no longer has the *desire* to do so. But this is of course not an option for the expressivist for whom moral judgments are understood as expressions of non-cognitive attitudes.

Another reply is to accept that moral judgments *normally* are the expressions of non-cognitive attitudes, but that in the case of severe depression, this is no longer the case. What the depressed agent expresses when she says that ϕ is good is rather her belief that she would normally have positive attitudes towards ϕ . But this is also not an option for the expressivist. The judgment “ ϕ is good” would on this interpretation still be an expression of a belief. Moreover, this interpretation would fit far better with subjectivism than with expressivism. What the depressed does is to *report* the feelings she normally has towards ϕ rather than to express an attitude towards ϕ .

But perhaps the expressivist can explain this problem by developing a more subtle account of which states of mind that are expressed in a moral judgment. As we will see in chapter 4, the expressivist need not say that moral judgments are the expression of first order attitudes. Rather, she can argue that moral judgments are the expression of higher order attitudes. Can this help the expressivist out? Let us say that the depressed makes the moral judgment that she ought to give money to charity. A simple minded expressivist will interpret this claim as an expression of a desire to give to charity. But since the depressed has no motivation to give to charity, it seems implausible to ascribe this desire to her. A subtler expressivist, however, will argue that what she really is expressing is a

¹⁶ Different versions of WMI will have different clauses to cope with counterexamples like Stocker’s. According to Michael Smith’s version of WMI (as mentioned in section 2.1), for instance, these counterexamples will fall under the “absent practical irrationality” clause.

second order desire. She desires to desire to give to charity. And she can have the second order desire while lacking the first order desire. So the expressivist can maintain her claim that moral judgments are the expression of non-cognitive states, even when someone lacks a first order desire to act on her judgment.

But on this new interpretation the expressivist faces a dilemma. A second order desire to ϕ is either necessarily or contingently connected to the agent's motivation to ϕ . If there is a necessary connection between the second order desire and the motivation to ϕ , the expressivist will have given no explanation of how the depressed can be making a moral judgment. She has *a desire to be motivated* to give to charity, but *no motivation* to give to charity. The expressivist may want to accommodate this by claiming that the connection between second order desires to ϕ , and motivation to ϕ is merely contingent. But if moral judgments are the expressions of second order desires, and second order desires are merely contingently connected to the motivation to act on moral judgments, the expressivist must give up on internalism all together. Either way, the expressivist cannot accommodate WMI.

Perhaps the most promising reply on behalf of the expressivist would be to say that the depressed person's judgment that it is right to ϕ is the expression of an emotion, but because of the depression, this emotion is no longer motivating. The problem with this answer is that it will be very difficult to explain which kind of emotions are supposed to be expressed in cases like these. Any kind of positive emotions towards ϕ seems to involve at least some motivational force. Moreover, a better description of the depressed person who says that ϕ is good but feels no motivation towards ϕ will accommodate the *absence* of positive feelings towards ϕ . So the best interpretation of what the depressed person expresses seems to be that it is a belief. But in that case the expressivist has to deny that the depressed person is making a genuine moral judgment.

We can now summarize the argument. According to WMI, a severely depressed person who has no motivation to do what she think is right could still make a moral judgment. This is because the depressed person will fall under the "motivational disorder clause" Expressivists hold that moral judgments express non-cognitive attitudes; if a judgment expresses a belief it cannot be a moral judgment. In order to accommodate the intuition, captured in WMI, that the depressed person is making a genuine moral judgment, even though she is not

motivated, the expressivist must make two claims. First, that the depressed person is expressing a non-cognitive attitude. Second, that this non-cognitive attitude, in the case of the depressed person, does not entail any motivation. But this, I have argued, cannot be done. So the expressivist must deny that the depressed person is making a genuine judgment. This means that expressivism and WMI are incompatible. According to WMI, the depressed person is making a genuine moral judgment. According to expressivism she is not. Therefore, only the strong and non-defeasible version of internalism, SMI, is open for the expressivist.¹⁷ This is problematic for the expressivist. WMI is a far more plausible version of internalism than SMI. So even though internalism in the weak form is an attractive position, this does not count in the expressivists favour. Indeed, the attractiveness of WMI gives cognitivism an advantage. Stocker's examples show that weakness of the will in the strong sense that I have discussed, is far more easily explained on the assumption that moral judgments express beliefs.

However, if moral judgments express beliefs, and if moral judgments motivate (except in the cases of motivational disorders that affect the agent more generally), beliefs must be able to motivate. This claim is denied by the expressivist, who is committed to the Humean Theory of Motivation (HTM), according to which a belief alone cannot motivate. HTM is the subject of the next chapter.

¹⁷ Allan Gibbard accepts this as a consequence of his version of expressivism. "To think, for instance, the thing to do is to defy the bully who torments me is to plan to defy him. [...] My theory thus yields internalism in a strong form: if I think something now is the thing to do, then I do it." (Gibbard 2003: 153). I take this to be a serious drawback of Gibbard's position.

3. The Humean Theory of Motivation

- 3.1 The Humean Theory of Motivation
- 3.2 An Anti-Humean Argument from Moral Phenomenology
- 3.3 A Dispositional Conception of Desire and the Teleological Argument
- 3.4 Can Directions of Fit Differentiate Beliefs and Desires?
- 3.5 Motivated and Unmotivated Desires
- 3.6 Besires
- 3.7 Summary and Preview

3.1 The Humean Theory of Motivation

If the reasoning of the last chapter is on the right track, the expressivist has to claim that there is a necessary and non-defeasible connection between moral judgment and motivation. As we have seen, this is not the most plausible version of internalism. But even if one were to accept SMI, this will not be enough to provide the expressivist with a knock down argument against the cognitivist. All that any kind of motivational internalism can show is that moral judgments play an essentially practical role in our lives; there is a close tie between what we take to be good and the actions we perform. But a cognitivist may still accept this and argue that it is moral *beliefs* that play this practical role: If I believe that I ought to give to charity, this belief is what motivates me to do so. I do not need any other motivational source. Indeed, this way of reasoning corresponds with how we normally speak of moral obligations: “Although I don’t desire to give any money to charity I believe that I should.” If motivational internalism is going to provide an argument for expressivism, it has to be combined with the view that beliefs alone cannot motivate. This is the Humean theory of motivation (HTM):¹⁸ Blackburn uses this theory of motivation to justify his expressivism:

“There is an insuperable obstacle to keeping ethics under the rule of Apollo. Suppose we think ethics is entirely exhausted by our beliefs. What then? Even the most magnetic star does not attract everyone. Beliefs do not normally explain actions: it takes in addition a desire or concern, a caring for whatever the belief describes [...] [A]s Hume saw, somewhere there will always have to be a practical, dynamic state: a concern or stance or attitude involved in translating a belief into action.” (Blackburn 1998: 90-91, see also 1984: 187-189)

¹⁸ And not “Hume’s theory of motivation” since the following sections will draw on arguments inspired by Hume, but not necessarily views Hume himself held.

Blackburn does not give an argument in favour of HTM; he just presupposes it. But as HTM is a controversial, although popular, position in moral psychology, we need some reason to adopt it. In this chapter, I will focus on Michael Smith's highly influential argument for HTM (Smith 1994a, ch 4). Smith gives the following definition of the Humean theory of motivation:

HTM: R at t constitutes a motivating reason of agent A to ϕ iff there is some ψ such that R at t consists of an appropriately related desire of A to ψ and a belief that where she to ϕ she would ψ (Smith 1994a: 92).

In order to have a motivating reason we thus need an appropriate desire in combination with a relevant means-end belief. Belief and desires are here taken to be distinct states of mind; for any belief B and desire D it is always possible to have B without having D, and *vice versa*. One of the attractions of this theory is that motivating reasons provide us with a straightforward explanation of action. I have a desire to quench my thirst. Together with my belief that I will quench my thirst by drinking a glass of water, this is enough to explain why I am drinking a glass of water. Both the belief and the desire are necessary in order to give an explanation but none of them are sufficient alone. If I am not thirsty, my belief that I will quench my thirst by drinking, does not give an explanation. Nor does my desire to quench my thirst amount to an explanation of my behaviour unless it is combined with the belief that I can do so by drinking a glass of water. Now, the claim that I had motivating reasons to drink a glass of water may be understood as a *causal* claim; my motivating reason to drink a glass of water provides a causal explanation of my action. (Davidson 1963). But Smith argues that a Humean need not make this claim. The basic explanatory concept that HTM provides is *teleological*. If I drink a glass of water because I have a motivating reason to do so, my action will have a teleological explanation; my action is intelligible in terms of the pursuit of a goal. Whether this teleological explanation is also a causal explanation is not determined by HTM. What causes actions is an empirical question which cannot be answered a priori.

3.2 An Anti-Humean Argument from Moral Phenomenology

According to HTM some relevant desire is needed in order to give the agent a motivating reason. But this does not fit well with the phenomenology of moral experience. We often experience that we have no desire at all to do what we take to be morally right. Nevertheless, we are motivated to do what we take to be morally right. Mark Platts offers the following argument:

“We perform many intentional actions in [our moral life] that we apparently do not desire to perform. A better description of such cases appears to be that we perform them because we think them desirable. The difficulty of much of moral life then emerges as consequence of the apparent fact that desiring something and thinking it desirable are both distinct and independent” (Platts 1979: 256, quoted from Smith 1994a: 110)¹⁹

As we are motivated to perform actions that we take to be morally right also in cases where we cannot, by introspection, find any desire to do so, it seems that the motivational job is being done by our *belief* that the action is morally right. Therefore, a desire does not seem to be a necessary condition for moral motivation. As Platts points out, the Humean might argue that to take something to be right, or desirable, *is* to desire it. But this does not seem to be true to the phenomenology of moral experience.

Smith answers this objection by arguing that it relies on a false conception of desires. The conception of desire that Platts argues against is a *phenomenological* conception: “the view that desires, like sensations, are states that have a certain phenomenological content” (Smith 1994a: 105). To be in pain, we presumably need to have a certain phenomenological sensation.²⁰ According to this view, the same goes for desires. A desire is state with a certain phenomenological feel.

Smith argues that this conception should be rejected because it implies an implausible epistemology of desire.²¹ We do seem able to be mistaken about our own desires. We think we have desires that we do not have, and we might have

¹⁹ For a similar argument against HTM, see Schafer-Landau (2003: 123).

²⁰ This is not uncontroversial, but it is difficult to find a better example of a phenomenological concept.

²¹ Smith also argues that the phenomenological conception cannot provide us with an explanation of why desires have a propositional content. I will leave the discussion of this argument aside in order to focus on Smith’s argument from the epistemology of desires.

desires which we are unaware of. Smith gives two examples of mistakes of this sort. Each day John buys newspapers at a certain newspaper stand on his way. He goes a longer route to do so instead of buying the newspaper at the stand which is on the direct route. The only difference between the newspaper stands is that the first has mirrors so placed that John sees his reflection when he buys his newspaper. John himself would deny that he has a desire to look in the mirror. But, as Smith describes the example, if the newspaper stand removes the mirrors, John's preference for the first newspaper stand would disappear. Smith argues that the most plausible explanation of John's behaviour is to say that he in fact has a desire to look in the mirror, but that he is unaware of this desire. In Smith's second example John believes that he has a desire to become a great musician. His mother has always wanted him to be a musician and John admits that he has a strong desire not to upset her. But he denies that this has anything to do with his desire to become a musician. When she dies, however, John loses all interest in music and gives up his musical ambitions. Smith takes this to mean that John thought he had a desire which he in fact never had. Therefore, he claims, we can have desire without being in state with a certain phenomenological "feel."

3.3 A Dispositional Conception of Desire and the Teleological Argument

Because of the problems facing the phenomenological conception we need an alternative conception of desires. Smith's starting point for providing this alternative is the thought that belief and desires have different directions of fit. This metaphor was originally developed by Elisabeth Anscombe (1957). Anscombe describes two ways of using a shopping list. First of all it can be used to direct actions. It tells the shopper what to do. But a detective who wants to record what the shopper buys can also make a list. In this case the list describes the shopper's purchases. In the former case the world should fit the desires; the shopping chart should be filled in accordance with the list. In the latter case the belief should fit the world. The detective's list should accord with what the shopper buys. The idea is that desires have a mind-to-world direction of fit. They aim to effect changes in world. Beliefs on the other hand, have a world-to-mind

direction of fit. They aim to describe or represent the world. Mark Platts gives the following account of how beliefs and desire differ in their directions of fit.

“Beliefs aim at the true, and their being true is their fitting with the world; falsity is decisive failing in a belief, and false beliefs should be discarded; beliefs aim to fit with the world, not *vice versa*. Desires aim at their realisation, and their realisation is the world fitting with them; the fact that the indicative content of a desire is not realised in the world is not yet a failing in the desire, and not yet a failing to discard the desire; the world, crudely, should be changed to fit with our desires, not *vice versa*.” (Platts 1979: 256-257, quoted from Smith 1994a: 112)

This is highly metaphorical, so Smith attempts to define desires in terms of their functional role. To have a desire, Smith argues, is to be in a complex dispositional state, a state that disposes one to act in certain ways. How we are disposed to act depends on the circumstances we are in, including our beliefs and our other desires. This dispositional concept of desire enables Smith to answer the difficulties that undermined the phenomenological concept. It improves on the epistemology of desires because it allows that agents are fallible about their desires. The claim that desires are states that disposes us to act does not entail that the agent must know his own desires. Moreover, the dispositional conception can also account for why desires sometimes do have a phenomenological content. Although all desires are dispositions to behave in certain ways, some desires are dispositions to produce feelings, and some are not.

Most importantly, the dispositional conception of desires enables Smith to cash out the direction of fit metaphor in terms of the different functions of belief and desires. Let us call this analysis FUNC

FUNC: [The difference between beliefs and desires] amounts to a difference in the counterfactual dependence of a belief and a desire that *p* on a perception that *not p*: roughly, a belief *p* is a state that tends to go out of existence in the presence of a perception with the content of *not-p*, whereas a desire that *p* is a state that tends to endure, disposing the subject in that state to bring about that *p*. (Smith 1994a:115)

Now Smith uses this functional conception of directions of fit to give an argument in favour of the Humean Theory of Motivation (HTM.) Following

Wallace (1990:359), I will call this the teleological argument. The argument consists of three steps:

- a) Having a motivating reason is, *inter alia*, to have a goal
- b) Having a goal is being in a state with which the world must fit
- c) Being in a state with which the world must fit is desiring.

Therefore having a motivating reason is, *inter alia*, desiring.

Smith argues that none of these steps can be denied. a) is, according to Smith, a conceptual truth; “we understand what it is to have motivating reason in part precisely by thinking of her as having some goal” (Smith 1994a: 116). b) seems straightforward given the functional account of desires of fit. A goal is not something that tends to go out of existence by the perception that the goal is not realised. Having a goal thus disposes us to effect changes in the world, rather than to change our goal to fit to the world. But does c) follow from a) and b)? It may be argued that desires are not the only mental states to which the world must fit. Hopes, wishes, intentions all share this direction of fit, Smith claims.²² Nevertheless it is not useful to label all these different mental states as desires. This is, however, no substantial objection. The Humean can just define the class of mental states with this direction of fit as *pro-attitudes* (as opposed to *beliefs*, which have the contrary direction of fit.) Now, the Humean might simply modify her claim and say that to have a motivating reason is, *inter alia*, to have a *pro-attitude*. As I will discuss in the next chapter, this is important for the quasi-realist project. The quasi-realist needs to hold on to the Humean view that a belief alone cannot motivate. On the other hand she will be reluctant to say that moral sentences express desires, narrowly understood.

3.4 Can Directions of Fit Differentiate Beliefs and Desires?

Smith’s functional analysis of beliefs and desires does much of the work in the teleological argument, so let us look at this more closely. Desires and beliefs have different functional roles. A belief that *p* tends to go out of existence in the

²² As I will discuss in the next section, the case of hope is somewhat problematic.

presence of the perception with the content of not- p , whereas a desire that p tends to endure. Adopting the terminology of Copp and Sobel (2001), let us call the belief or desire that p the background state of the agent. The background state is to be distinguished from “the perception with the content of not- p .” Let us call this the introduced state. As many commentators have pointed out it is not at all clear how we are to understand the introduced state. If it is to be understood as a sensory perception it would be too narrow. If I read an article about the extinction of the great auk, we would not say that I had a sensory perception that the great auk is extinct. Still, this would dispose me to give up my belief that there are great auks in the world. Another possible way of interpreting the introduced state is to understand it as a belief. Hence, when I read about the extinction of the great auk, I form the belief that the great auk does not exist. But, as Copp and Sobel points out, this would make the analysis circular. An analysis of what belief and desires are cannot presuppose the notion of belief. However, it might be objected that Smith is not offering an analysis of desires and beliefs. What he is aiming at is rather to differentiate these two states. But if this is the case, it would still be problematic to take “perception” as another word for belief. If the difference between desires and beliefs is characterized by reference to what is being differentiated, i.e. a belief, Smith’s account would be uninformative.²³ We therefore need a different way of understanding the nature of the introduced state.

The reading which Smith himself endorses (Humberstone 1992: 64, note 10) takes “the perception with the content that p ” to mean that it “appears” or “seems” to the agent that p . This is not the same as the belief not- p . As I walk down the street it might seem to me that my grandfather is walking towards me. But given that he is dead, I do not believe that he is walking towards me. This way of understanding the introduced state avoids circularity. But it leads to another problem.

Given this new interpretation of the introduced state, we can reformulate Smith’s functional analysis:

²³ Copp and Sobel (2001) argue that Smith offers an analysis and the inclusion of belief in the introduced state makes this analysis circular. Humberstone (1992) claims that Smith does not offer an analysis and that the inclusion of belief is merely uninformative.

FUNC2: A belief p is a state that tends to go out of existence when it seems or appears to the agent that not- p , whereas a desire that p is a state that tends to endure, disposing the subject in that state to bring about that p .

But this analysis is vulnerable to counterexamples. The fact that it appears to an agent that not- p does not imply that the agent tends to give up his belief. Copp and Sobel give the example of someone who is driving on a hot pavement. It appears to him that there are bubbles on the road. But an “experienced driver has no tendency at all to cease thinking that the road is dry when confronted with such an appearance.” (Copp and Sobel 2001: 47). Another example would be familiar cases of wishful thinking. The doctor may tell me that I only have a week left to live, but this does not mean that I will give up my belief that I will live. According to Smith’s account, this attitude should be categorized as desire, rather than a belief. But this seems strange. To be subject to wishful thinking is often considered to be irrational, or blameworthy. But if wishful thinking was just a species of desires, there is nothing irrational about it. My desire to live is perfectly in order, whereas my belief that I will live might be irrational.

These counterexamples have dealt with beliefs. But the similar point can be made about desires. I might have a desire that p which goes out of existence when it seems to me that not- p . Copp and Sobel give the example of Sue, the fair weather fan. She desires that the 49’ers do well, but when they are not doing well this desire disappears. This does not mean that she desires that whichever team happens to well does well. The content of her desire is that 49’ers to do well. But her desire goes out of existence when she learns that they are losing. Still, it does not seem helpful to say that her mental state really is the belief that 49’ers will win.

Smith also incorporates pro-attitudes as states with the same direction of fit as desires. But there are other pro-attitudes as well, which go out of existence when it appears to the agent that they are not realised. As G.F Schueler (1991) has argued, hopes provide us with an example of this. I cannot continue to hope that Manchester United will win the Champions League final when I see on television that they have just lost. I might of course wish that this was not the case. But I cannot continue to hope that they will win. But this does not mean that hopes are beliefs.

As a reply to these counterexamples Smith might argue that taken one by one, there will be cases which do not fit his analysis. But most of our beliefs and desires do have the constituting dispositions. It is thus the class of beliefs and the class of desires that has a distinct functional role. But given that there are beliefs and desires which do not have the direction of fit that is characteristic for the *class* to which they belong, we cannot use the direction of fit to distinguish between *individual* cases of belief and desires. From the fact that a mental state with the content that *p* tends to endure when it appears to the agent that not *p*, we cannot infer that it is a desire. Similarly, from the fact that a mental state with the content that *p* tends to go out of existence when it appears to the agent that not *p*, we cannot infer that it is belief. Thus, Smith's direction of fit account fails to provide a satisfying way of differentiating individual beliefs from individual desires. (Schafer-Landau 2003: 35)

These counterexamples aim at destroying Smith's argument before it gets started. If the functional analysis fails, talk of directions of fit will remain highly metaphorical and of little use to the advocates of HTM. But it might be the case that it is possible to amend the failures of FUNC2. Moreover, the direction of fit metaphor does seem to capture something intuitively correct about the distinction between desires and beliefs. For the rest of this chapter I will assume, for the sake of the argument, that some functional analysis of direction of fit can be given. Only if we make this assumption, Smith's teleological argument is successful: Having a motivational reason is desiring.

This weakens the expressivist's case. Let us recapitulate the dialectical situation. The expressivist must argue for a combination of internalism and HTM in order to give an argument for the claim that moral judgements express non-cognitive attitudes and not beliefs. The last section, however, showed that there are some grave difficulties with Smith's teleological argument for HTM.

But if we assume that these difficulties can be overcome, does this provide the expressivist with an argument that moral judgments must be the expressions of non-cognitive attitudes? This is not clear. I will now consider two ways of *accepting* that motivation necessarily involves desire, while at the same time denying that beliefs cannot motivate. If one of these two strategies is successful, it would further weaken the case for the expressivist.

3.5 Motivated and Unmotivated Desires

In *The Possibility of Altruism*, Thomas Nagel makes a distinction between unmotivated and motivated desires (Nagel 1970: 29). Unmotivated desires are desires like thirst and hunger and some of our emotions. These desires simply assail us and thus motivate to action. Motivated desires, on the other hand, are not sources of motivation, but should rather be understood as products of some consideration, or evaluation.²⁴ Duties, values and commitments fall into this category. I have the desire to be a good student, but this desire is not the same kind of desire as my desire to quench my thirst. In the former case I find that I have reasons to be a good student, or I have an evaluative belief that I ought to be a good student. According to Nagel, it is not the case that I make the evaluation that I ought to be a good student because of my desire. It is rather the other way around: my desire to be a good student is a consequence of my evaluation that I ought to be a good student. The difference between motivated and unmotivated desires is that the former can be given a rationalizing explanation.²⁵ Desires, in combination with the relevant means-ends beliefs provide us with a teleological explanation. Given a desire and the relevant means-end belief, we can explain an agent's behaviour in terms of her goals. But the agent is not necessarily aware of her own desires. As we saw in paragraph 3.2, this is in accordance with Smith's dispositional concept of desires. Nor must the agent accept her own desires as something desirable. A drug addict may desire her drug and thus be motivated to take it. In Smith's terminology we therefore speak of the drug addict's goal of taking the drug. But this does not mean that the addict finds her desire desirable. A rationalizing explanation, on the other hand, is an explanation that refers to the reasons why the agent takes something to be desirable. A rationalizing explanation of my desire to be a good student will thus cite the reasons I take in favour of being a good student. The cognitivist will say that taking something to be desirable is to have an evaluative belief. When I have a motivated desire I have this desire *because* of my evaluative belief. Thus, according to Nagel, I have the motivated desire to be a good student because I have the evaluative belief that I ought to be a good student. We can now see why Nagel accepts that motivation

²⁴ Motivated and unmotivated desires should however not be taken as a comprehensive typology of desires. There seem to be desires that fit into neither of these categories. See Wallace (1990: 363).

²⁵ For this way explaining Nagel's notion of motivated desires, see Wallace (1990.)

necessarily involves having a desire. He can grant this, but still maintain that some of our desires are produced by our evaluative beliefs. So although desires are necessary in order to motivate, desires might be generated by evaluative beliefs.²⁶ This means that the Humean picture of motivation is turned up side down. Far from being passive and inert, some of our beliefs are, according to Nagel, able to play an active and dominating role, namely by causing desires.

There is, however, a problem with this suggestion. This concerns the connection between evaluative beliefs and motivated desires. This connection can be understood in two ways. The desires are either caused by the evaluative beliefs or motivated by them. If the relation is causal the Humean need not to deny it. A merely causal relationship between evaluative beliefs and desires does nothing to challenge the claim that a desire and a means-end belief are necessary for motivation. (Dancy 2000: 82-84). So what Nagel needs is that beliefs *motivate* desires. But if we accept Smith's teleological argument, this cannot be. According to Smith no beliefs can motivate because they have the wrong direction of fit. If desires are to be motivated by beliefs, we need to explain this by referring to a motivating reason. But at this point, the Humean may simply rerun the teleological argument: a) Having a motivating reason is, *inter alia*, to have a goal, b) Having a goal is being in a state with which the world must fit, c) Being in a state with which the world must fit is desiring. Thus, if the Humean is to accept the distinction between motivated and unmotivated desires, she will claim that the motivated desires must be motivated by another desire, and not an evaluative belief. This means that an evaluative belief cannot play the active role that Nagel ascribes to it. So if we accept Smith's teleological argument, Nagel's distinction between motivated and unmotivated desires does not provide the cognitivist with an argument against HTM.

3.6 Besires

Another way of arguing against HTM is to claim that some of our evaluative beliefs, moral beliefs in particular, have *both* directions of fit. (McDowell: 1978) On the one hand they have a mind-to-world direction of fit

²⁶ There is an ambiguity about what Nagel meant by "motivated desires." (Dancy 1993: 8) On one interpretation motivated desires have an independent existence insofar that they are caused by evaluative beliefs. On the second interpretation "desires" is just another word we use to indicate that the agent is motivated. I have adopted the first interpretation.

because they dispose us to action. On the other hand they have a world-to-mind direction of fit because they tend to go out of existence if we perceive that our evaluative belief is not true. Take the evaluative belief that it is good to be honest. According to this Anti-Humean position, the evaluative belief will motivate me to tell the truth. But if I discovered that it in fact was not good to be honest - perhaps I come to see that kindness is much more important, and that honesty often implies unkindness - then my belief that honesty is good would disappear. If evaluative beliefs have both directions of fit, it follows that the Anti-Humean must postulate a third mental state, in addition to the typical cases of belief and desires. This third state is what J. E. J Altham (1986) has named “besires.” These mental states will share the both the dispositions we associate with beliefs, as well as the dispositions we associate with desires. A person who has an evaluative belief will therefore be in a state that is both belief-like and desire-like. No additional desire would be needed to motivate the agent.

But the Humean will claim that beliefs and desires are distinct existences; it is always possible to be in a belief-like state without being in an appropriately related desire-like state and *vice versa*. To claim that there are besires amounts to postulating a single, unitary state that is both desire-like and belief-like. But if beliefs and desires can always modally be pulled apart this unitary state is impossible. Michael Smith (1994a: 119-120) argues as follows. If the belief that I ought to ϕ was part of some single state, it will be impossible for me to believe I ought to ϕ and yet fail to be motivated to ϕ . But this commits the besire theorist to a far too strong version of motivational internalism. It is possible to believe that I ought to ϕ without being motivated to ϕ . Recall the two versions of internalism that we encountered in the previous chapter:

SMI: If an agent judges it is right for her to ϕ in circumstances C, then she is motivated to ϕ in C.

WMI: If an agent judges it is right for her to ϕ in circumstances C, then she is motivated to ϕ in C, except in cases where the agent is suffering from motivational disorders that affect her more generally.

SMI is the non-defeasible variant of Internalism, WMI is defeasible, because “motivational disorders” is meant to cover cases of extreme weakness of will, *accedia* and depression where the agent finds no motivation to do what she believes is right. Smith’s point is that the postulation of *besires* implies the strong and non-defeasible version of internalism. But, as discussed in the last chapter, Stocker’s counterexamples make SMI problematic.

We are now in an odd dialectical situation. In the last chapter I argued that motivational internalism is plausible, but only in its weak form. This version of internalism is not available to the expressivist because she has to claim that making a moral judgment involves some kind of emotional endorsement. And this is exactly what is lacking in the cases of depression and *accedia*. So the weak form of internalism support some form of cognitivism, because Stocker’s examples imply that moral beliefs and the appropriately related desires can come apart. But since moral judgments normally motivate, the cognitivist has to argue against the teleological argument for HTM. As we have seen in section 3.4, the functional analysis on which the teleological argument rests faces some grave difficulties. But if we assume that these problems can be amended, the most promising way of arguing against HTM is by claiming that moral beliefs are “*besires*.” But “*besires*” are mental states that necessarily involve both beliefs and desires. The cognitivist will therefore have great difficulties in accommodating Stocker’s examples. So if the postulation of *besires* is the best answer to the teleological argument, the cognitivist loses his dialectical advantage over the expressivist. Whether we accept or deny HTM, we seem nevertheless forced to accept an implausible version of internalism.

As far as I can see, there are three different solutions to this problem. We can accept HTM and argue for SMI. This is the expressivist solution. We can deny HTM and argue for SMI. This is the *besire* theorist’s solution. Or, finally, we can accept HTM but deny internalism altogether.²⁷ This is the externalist solution. The problem for the expressivist is that she needs *both* a plausible version of internalism and the Humean theory of motivation in order to give an

²⁷ However, if we take the problems facing Smith’s teleological argument to be unsolvable, this leaves the possibility open for denying HTM, without postulating the existence of *besires*.

argument for her position based on the motivational features of moral judgements.²⁸ And this, I have argued, cannot be given.

3.7 Summary and Preview

We can now return to the moral problem, which I discussed in the introduction. There I stated three independently plausible claims which together form an apparently inconsistent triad. The first claim was that moral judgments aim at objectivity, the second was that moral judgments necessarily motivate and the third was the Humean theory of motivation. Drawing on the discussion of the last two chapters, we can now restate the moral problem in a more precise manner:

1) Moral judgments a) aim at objectivity, so b) moral judgments are beliefs.

2) (WMI) If an agent judges it is right for her to ϕ in circumstances C, then she is motivated to ϕ in C, except in cases where the agent is suffering from motivational disorders that affect her more generally.

3) (HTM) R at t constitutes a motivating reason of agent A to ϕ iff there is some ψ such that R at t consists of an appropriately related desire of A to ψ and a belief that where she to ϕ she would ψ .

2) and 3) were supposed to provide us with a good reason to deny 1b). However, as we have seen, the argument from motivational internalism and the Humean theory of motivation is not as straightforward as one might wish. None of the positions I have discussed in the last two chapters can accommodate both WMI and HTM. Both expressivists and desire theorists have great difficulties in accounting for the truth of WMI, and externalists deny it altogether. I will therefore turn my attention to moral objectivity. If Blackburn can show that the moral objectivity can be saved without invoking moral facts or moral beliefs, quasi-realism will still be a very attractive theory. Since none of the competing

²⁸ There are of course other ways of arguing for expressivism, especially in its quasi-realist trapping. But in this and the previous chapter I was only concerned with the argument from moral motivation.

theories can account for WMI, quasi-realism will have a dialectical advantage because it promises to accommodate moral objectivity without depending on any heavy-duty metaphysics. To evaluate Blackburn's attempts to do this will be the aim of the next three chapters.

4 Moral Attitudes and the Demarcation Challenge

4.1 Attitudes and Normativity

4.2 Higher Order Attitudes

4.3 Stable Attitudes

4.4 Prescriptive Attitudes

4.5 Moral Authority and Justification

4.1 Attitudes and Normativity

The previous chapters focused on the motivational aspect of moral judgments. According to Blackburn, moral judgments are the expressions of non-cognitive attitudes, such as positive or negative feelings, desires or aversions. While it is uncontroversial that these states of mind are able to motivate us, not all of our desires or feelings can count as *moral* attitudes. Normative judgments involve a claim of authority, but it is not the case that any old emotion or desire has this authority on us. Although desires motivate us they do not necessarily give us normative reasons to act. What we desire is not the same as what we value. An addict might well have a desire to take her drug, but does not thereby think that she ought to take the drug. I might be angry with someone for criticizing me and desire to give a scorching reply. But if I admit that the criticism is justified, I might well think that I have no reason to be angry, and that my desire is unjustified. Being in an emotional state is therefore not the same as making an evaluative judgment. There is a gap between our actual emotional responses and the judgments that we think are justified. In order to show how we can make sense of the claim that moral judgments express attitudes, or sentiments, the quasi-realist has to give an account of which kinds of attitudes that are expressed in a moral judgment. She must provide criteria by which we can demarcate the proper subset of our attitudes which are moral attitudes from those of our attitudes that are not. This is the demarcation challenge.

Blackburn gives the following account of the connection between simple likes and dislikes and full-blooded moral judgments. It is worth quoting in full.

“We should think in terms of a staircase of practical and emotional ascent. At the bottom are simple preferences, likes and dislikes, more insistent is a basic hostility to some kind of action or character or situation: a primitive aversion to it, or a disposition to be disgusted by it, or to hold it in contempt, or to be angered by it, or to avoid it. We can then ascend to reactions to such

reactions. Suppose you became angry at someone's behaviour. I may become angry at you for being angry and express this by saying it is none of your business. Perhaps it was a private matter. At any rate, it is not a moral issue. Suppose, on the other hand, I share your anger or feel "at one" with you for so reacting. It may stop there but I may also feel strongly disposed to encourage others to share the same anger. By then I am clearly treating the matter as one of public concern, something like a moral issue. I have come to regard the sentiment as legitimate. Going up another step, the sentiment may even become compulsory in my eyes, meaning that I am prepared to express hostility to those who do not share it. Going up another level, I may also think that this hostility is compulsory, and be prepared to come into conflict with those who, while themselves concerned at what was done, tolerate those who not care about it, I shall be regarding dissent as beyond the pale, unthinkable. This should all be seen as an ascending staircase, a spiral of emotional identifications and demands. The staircase gives us a scale between pure preference, on the one hand, and attitudes with all the flavor of ethical commitment on the other." (Blackburn 1998: 9)

This demarcation of moral attitudes relies on three factors. First, the attitudes characterized as moral are higher order attitudes. Secondly, we identify with these attitudes more strongly than with mere preferences. Thirdly, the moral attitudes are characterized by an emotional intensity that is typically directed towards others. In the cases of "full flavoured" ethical commitments, we demand that others share our higher order attitudes, and we express hostility against those who do not. I will discuss these three factors in turn.

4.2 Higher Order Attitudes

On Blackburn's account, moral judgments should not be understood as the expressions of simple emotional reactions such as hostility towards some act, or approval towards some other act. Rather the moral judgment consists in forming higher order attitudes towards the initial attitude. I might have a strong aversion against hypocrisy but I might also form the higher order attitude toward my aversion against hypocrisy; I approve of my aversion. Higher order attitudes like these may be formulated in terms of norms. When I claim that hypocrisy is wrong what I do is to accept some norms which endorse my first order attitude, i.e. my aversion towards hypocrisy. In order for the present position not to collapse into cognitivism, the higher order attitudes cannot be understood in terms of beliefs. The higher order attitudes must themselves be understood as non-cognitive attitudes towards my first order attitude. This might seem as plausible explanation

of the addict's judgment that she ought not to take her drug although she desires it. She does not value what she desires because valuing is to be understood as a second order desire. The drug addict does not desire to desire to take her drug (Frankfurt 1971; Lewis 1989).²⁹ So except in cases of weakness of will, our second order attitudes govern our first order attitudes.

That an attitude is a higher order attitude is, however, not enough to demarcate it as a moral attitude. Let us imagine a conflict between a first order and a second order desire. Jones desires to be honest. On the other hand he has a desire not to desire to be honest. But his second order desire does not necessarily have any more of the features of a moral attitude. Perhaps Jones desires not to desire to be honest because he knows that his life would be a lot easier if he did not have the desire to be honest. In this case his desire to be honest conflicts with his concern for his own well being. In this case, Jones' second order desire is not a morally informed desire at all but rather a desire based on a prudential concern. Second order attitudes are not moral attitudes just because they are second order; we need a supplementary way of demarcating second order attitudes that will count as moral attitudes from those which cannot. I will return to Blackburn's preferred way of doing this below.

Conflicts between first and second order supports the claim that second order attitudes does not suffice to account for the authority of moral attitudes. Why should we revise our first order attitude in the light of our second order attitudes and not the other way around? What is it about second order attitudes that give them a special authority? According to Harry Frankfurt (1971) the answer to this question is that the agent identifies more intimately with second order attitudes.³⁰ So they are the agent's own in a way that first order attitudes are not. This is the reason why second order attitudes are taken to be a special source of authority in cases of conflict between first and second order attitudes. But as John Watson (1975) has pointed out, it is doubtful whether second order attitudes really have this authority. As second order attitudes themselves are only attitudes it is unclear why they should be any more closely related to the agent than the first

²⁹ Although David Lewis (1989) is the most outspoken proponent of understanding values as second order desires, I will not deal in depth with his reductive analysis of values. This is because he develops a dispositional account of values which differ from Blackburn quasi-realism.

³⁰ Frankfurt speaks of second order desires, but I will use "attitudes" for the sake of consistency with Blackburn's vocabulary.

order attitudes. The problem with ascribing normative authority to attitudes was that we often do not find anything desirable in our pro-attitudes. So, moral judgments cannot be the expressions of simple attitudes. But we can now see that the same problem arises if we ascribe normative authority to second order attitudes. Just as with first order attitudes, I may fail to find anything desirable in my second order attitudes. For any second order attitude, there might be a conflicting third order attitude. It is always possible that I do not find my second order attitude desirable; I may desire to desire ϕ , yet still fail to value ϕ . Perhaps I do not have any desire to be wealthy. On the other hand I do have the desire to desire of being wealthy. If I had this first order desire, my financial situation would probably be better than it is. But of course it is quite conceivable that I have a third order desire not to have the second order desire. I might for example wish that I did not care about economical matters at all. In this case there is no reason why the second order judgment should have a privileged position. A first order attitude may conflict with a second order attitude, which may conflict with a third order attitude, which may conflict with a fourth order attitude, and so on. There is no non-arbitrary way of determining which level of higher order attitudes that should be expressed in a moral judgment. (Smith 2004: 371- 374)

Because of this difficulty, it would be tempting to align moral attitudes with the highest order attitude. But as Blackburn himself points out, this suggestion does not work either. People might regret having the values they do. We can imagine a virtuous person, who, in weak moments has a highest order desire not to be as virtuous as she is; she might wish that she was more relaxed, for instance. But this does not mean that she values being more relaxed. Her highest order desire should here be understood as more of a whim or temptation than a value. (Blackburn 1998: 66-67)

That higher order attitudes by themselves are incapable of providing the normativity of moral judgments, is perhaps most easily appreciated if we consider the differences between a first and a higher order attitude. The only difference is that the higher order attitude has the first order attitude as an object. But this difference is not enough to give the higher order attitude authority over the lower order attitude. This point becomes clear if we instead of focusing a conflict within a person focus on a conflict between two persons. We often form higher order attitudes towards other people's attitudes. Jones is angry with Smith. Smith

becomes angry with Jones because of Jones' anger. Smith's attitude is a higher order attitude, but this does not mean that Smith's anger at Jones anger should have authority over Jones anger with Smith. (Scanlon 1998:54; Kalderon 2005: 171)

4.3 Stable Attitudes

Something more is needed in order to locate the attitudes that are expressed in a moral judgment. Blackburn admits this, and makes the following suggestion.

“Desire to desire does not do either. [...] A better general account might locate the ethical in the springs of action with which the agent is most identified, and that in turn would be manifested by things like reluctance to change or reluctance to tolerate variation.” (Blackburn 2006a: 150)

These specifications are more direct attempts to describe what the higher order attitude account was meant to capture. As I argued in the last section, it does not follow that an agent will identify herself with an attitude just because it is a second order attitude. There might be second order attitudes by which an agent identifies herself, but also second order attitudes by which she does not. In the passage quoted above, Blackburn seems to acknowledge this by postulating that a moral attitude is an attitude by which the agent identifies herself. Moreover, this identification must exhibit some stability. These two features are closely related. Valerie Tiberius (2000) has developed this aspect of quasi-realism. She asks us to consider Jean Valjean, the main character from Victor Hugo's *Les Misérable* as an example of these two features of moral attitudes. Inspired by a Monsignor, the ex-convict Valjean promises that he will become a new man. Under a new identity he becomes a successful businessman and also a man devout to God. When he discovers that another man is accused of committing a crime he himself perpetrated did many years earlier, Valjean cannot let the other man unjustly suffer and reclaims his identity. According to Tiberius, Valjean's commitment to his ideals serves as an example of how tightly moral attitudes are related to our identity; to the persons we want to be. Our moral attitudes are of great importance to us, they enjoy a privileged role in our choices of action. So, as noted in chapter

two, we are normally motivated to act in accordance with our moral judgments. But we are motivated by a lot of things that we do not consider to be morally significant. Blackburn's and Tiberius' point here is that the attitudes we form faced with a situation or choice of actions are attitudes we approve of because they constitute the persons we want to be. The specifically moral attitudes, Blackburn suggests, are the attitudes by which we identify ourselves. The connection to the second requirement, that our moral attitudes must be stable, should now be clear. As Tiberius points out, the attitudes we approve of must be stable in order to compensate for the fluctuating nature of our motivations. Moreover, they must be stable in order to figure in the planning and deliberation of future actions. We rely on the attitudes we identify with not only for the choices we are presently confronted with, but also for future actions, and our way of life. If the attitudes in question were unstable they could not serve as an identifying feature. If we identified ourselves by something constantly changing, we would either be insincere or suffering from a fragmented personality.

These two features seem to capture an important difference between moral judgments and mere preferences, between desires and what is desirable. This can be seen if we consider how different kinds of desires manifest themselves in the choices we make. If I make a choice between spending my holiday in France or in Peru, or between eating a pizza or a hot dog for lunch, I do not thereby identify myself in terms of these choices. I opt for one of these alternatives because I feel like it, or because I only can afford travelling to France or buying a hot dog. Moreover, choices like these are mutually exclusive. Since I cannot both go to France and Peru at the same time, I choose France. If I, on the other hand, must choose between telling my girlfriend about an affair, or remaining silent on the matter, my choice involves questions about whom I want to be. I do not choose to tell her because I just feel like it, nor simply because I must choose one of the two possibilities (although this of course is true). The difference between these two kinds of choices is the difference between what Charles Taylor has called weak and strong evaluations. What characterizes strong evaluations, as opposed to weak evaluations, is that the choice can be described in a qualitative language and that they relate to the question of whom I want to be. I choose to tell my girlfriend because it would be the *honest* thing to do, or I remain silent because this is the *kindest* thing to do. In cases like this, our different desires "are judged as

belonging to qualitatively different modes of life: fragmented or integrated, alienated or free, saintly or merely human, courageous or pusillanimous and so on” (Taylor 1985:16).

There are, however, two interrelated difficulties with this way of demarcating moral attitudes. On the one hand it might be too strong. It can be argued that questions of personal identity and questions of morality should be separated. If this is the case, we might not need to identify with the attitudes in question in order to make a moral judgment, nor must the attitude expressed necessarily be a stable one. Imagine Jones, the shopkeeper, who endorses and identifies with a stable attitude of his; that of being stern. One day he catches a shoplifter red handed. Normally he would call the police at once, but this time he is struck by a rare feeling of compassion and lets the shoplifter go. Afterwards Jones regrets his action and acknowledges that it was out of character. Still, it seems plausible to assume that Jones made a moral judgment, although he did not act on a stable attitude of his.

On the other hand this criterion is too weak. Moral attitudes relate to other people in a way that questions of personal identity and commitment do not. Invoking the notions of identification and stability seems to be an appropriate way of characterizing our commitments and values. But moral judgments are not the same as expressions of what we value. Whereas our commitments and values relate to questions of how I live my life, moral judgments are directed towards others in a much stronger sense; they have a claim to universality built into them. I might have a variety of values and commitments by which I define myself as a person, but I do not necessarily want everybody else to share my commitments.

This point may have been somewhat obscured by the examples I have used to illustrate Blackburn’s identification-and-stability-account of moral attitudes. These examples have been of straightforward moral dilemmas such as telling a girlfriend about an affair, or Valjean’s choice of revealing his identity, even though he would lose his place in society. But by changing the examples, we can easily see that identification and stability are features that cover a wider area of human attitudes than the attitudes that are expressed in a moral judgment. A football fan will typically have a stable pro- attitude towards the club of his heart. Moreover, he might well identify himself in terms of being, say, a Tottenham fan. Indeed, his love for the club may very well be the most stable and unwavering

commitment of his life. He might even describe the question of supporting another club, (after another miserable season) in the qualitative language of strong evaluations. It would perhaps give him more pleasure to be a Middlesbrough fan, but this would be a cowardly and unfaithful thing to do. He does not want to identify himself with such treacherous desires as leaving his club for a more successful one. Arguably, there are lot of choices concerning questions of how to live my life and whom I want to be that are not moral choices. The question of which education I should choose, which kind of work I should apply for, which of my abilities I should cultivate all fall somewhere in between a simple ranking of preferences and a full blooded moral issue.

How and where to draw the line between questions of which kind of person we want to be, and which kind of life we want to lead on the one hand, and purely moral considerations on the other, is of course a delicate and substantial question in moral philosophy.³¹ But it would be a major drawback of quasi-realism if it were to rely on a controversial conflation of “the good life” and “the right thing to do.” Hence, stability and identification are insufficient to demarcate moral attitudes from non-moral commitments. Again, something more is needed: an account of how the moral attitudes relate to others.

4.4 Prescriptive Attitudes

Blackburn provides us with an account of this sort. As seen in the long quotation above, he strongly emphasizes the prescriptive aspect of moral attitudes. When we judge something to be morally good or obligatory we are not making this judgment merely on behalf of ourselves. Rather, we are demanding that everybody accepts our judgments. We are not indifferent to whether our audience accepts our claim. This aspect of moral judgments is clearly reminiscent of the work of Charles Stevenson. Stevenson (1937; 1944) argued that moral judgments, as opposed to the expression of beliefs, contain a do-so-as-well-aspect. By expressing approval or disapproval we want our interlocutors to share our attitude. We think that the emotional reactions we form when confronted with a situation

³¹ For an argument for separating these questions, see for instance Habermas (1991). The debate about whether morality should be separated from questions of personal identity is long and thorny, reaching from Kant’s *Moralität* vs. Hegel’s *Sittlichkeit* to the present debate between liberals and communitarians. This discussion falls outside of the scope of this thesis.

or action are reactions that ought to be shared by others. This do-so-as-well-aspect that Blackburn incorporates into his account of moral judgments seem to be a promising way of demarcating moral attitudes from mere desires, tastes and dislikes. If I express my hunger by saying: “I am hungry”, I do not thereby insist that others should share my desire to eat. If I, on the other hand, claim that torturing children is wrong, I will not tolerate dissent; I demand that you share my judgment. Moreover, it demarcates moral attitudes from non-moral commitments. Although the Tottenham fan has a stable pro-attitude towards his club, he does not want everybody else to share this commitment.

However, if this is all that is to be said about moral attitudes, quasi-realism faces the same objections that have been raised against emotivism, Stevenson’s (1937; 1944) brand of expressivism. This objection, most famously argued by Alasdair MacIntyre (1981), implies that moral judgments are a way of manipulating others. This is so because it construes moral judgments as containing a do-so-as-well aspect, a prescriptive element, without giving any justification for this claim. Epistemologically, emotivism denies that moral judgments can be justified. As a theory of the meaning of moral judgments, however, it maintains that moral judgments contain a demand to share the speaker’s emotions. But since no justification is offered, this means that moral judgments are vehicles of manipulation:

“Stevenson ... understood very clearly that saying “I disapprove of this; do so as well!” does not have the same force as saying “That is bad!” He noted that a kind of prestige attaches to the latter, which does not attach to the former. What he did note however – precisely because he viewed emotivism as a theory of meaning – is that the prestige derives from the fact that the use of “That is bad!” implies an appeal to an objective and impersonal standard in way which “I disapprove of this; do so as well does not!” (MacIntyre 1981: 19-20, quoted from Kalderon 2005: 32)

The do-so-as-well-aspect of moral attitudes is plausible enough. When we make moral judgments we lay pressure on others and we think that they ought to share our attitudes. But this characterization of moral attitudes needs to be supplemented with an account of the *justification* of this demand. If not, the quasi-realist account of moral judgments will have the same revisionary consequences as Stevenson’s emotivism.

4.5 Moral Authority and Justification

We are now in a position to summarize Blackburn's way of characterizing the mental states that are expressed in moral judgment. These mental states are typically complex emotional and motivational states, and involve higher order attitudes. These higher order attitudes are stable and are something with which the agent identifies herself. Moreover, the attitudes are directed towards others, thus manifesting a do-so-as-well aspect. However, to complete his characterization of moral attitudes, Blackburn needs to incorporate a notion of how these attitudes can be justified.

Blackburn claims that a quasi-realist has no problem accounting for this. Our evaluative language is not limited to simple expressions like "Boo for lying!", and we need not confine ourselves to commandments like "Don't lie!" We discuss and negotiate our moral attitudes; we rank them and compare them, accept and reject them, we are able to reason about attitudes and systemize them into coherent frameworks. The quasi-realist will even allow that opinions can be true or right. This is because "[a]sking whether a moral judgment is true or right is no more than asking whether to accept it. And asking this is asking which attitude or policy or stance to endorse" (Blackburn 1999: 214).³²

But given that the quasi-realist explains moral judgments in terms of attitude, and cannot rely on any heavy duty metaphysics of moral facts or properties, there is a common worry that the normative force of our moral judgments are dependent upon the attitudes we happen to have. After all, in virtue

³² By arguing this way Blackburn relies on minimalist or deflationist conception of truth. According to minimalism, saying that *p* is true is just to say that *p*. The truth predicate is a thin concept; by using it we are not committing ourselves to any heavy metaphysical or epistemological duties. Blackburn makes this point by describing what he call Ramsey's ladder: "Because of the minimalism we can have for free what look like a ladder of philosophical ascent: '*p*', '*it is true that p*', '*it is really and truly a fact that p*'..., for none of these terms, in Ramsey's view marks an addition to the original judgment. You can just as easily make the last judgment as the first - Ramsey's ladder is lying on the ground, horizontal." (Blackburn: 1998: 78) Armed with this notion of truth, the quasi-realist will have no difficulty in handling the concept of truth. On the minimalist notion "'Giving to charity is good' is true" simply means "Giving to charity is good". And "Giving to charity is good" can be interpreted as expressing some positive attitude to *X*. Blackburn applies the same strategy to our talk of moral facts. "It is a fact that lying is wrong", is simply a way of saying that lying is wrong. Since my focus in this thesis is on "the moral problem" and not on the semantic program of quasi-realism, I will not take issue with minimalism about truth, or Blackburn's use of this view. It is worth noting, however, that minimalist or deflationist theories of truth are controversial, and so is the question of whether a quasi-realist can make use of them. For the latter question, see for instance Smith (1994b), Horwich (1994) and Divers and Miller (1994).

of what is one attitude better, or more correct than another? This lack of objective and external standards seems to conflict with our experience of moral demands as something non-optional and categorical. So the quasi-realist must provide us with some notion of how standards of correctness can be applied to ethics. She must account for how some attitudes can be better than others, how we can criticize others for making moral errors, and how we can make sense of our own fallibility on ethical matters.

These issues are the subjects for the remaining two chapters. In the next chapter I will discuss Blackburn's answer to the objection that morality in some way is dependent on our attitudes. The last chapter will focus on how Blackburn can account for the possibility that our moral attitudes are wrong.

5. Mind-dependence and the Internal Reading.

5.1 Error and Mind-Dependence

5.2 The Internal Reading

5.3 The Internal Reading and Relativism

5.4 The Internal Reading and Generality

5.1 Error and Mind-Dependence

One of the seemingly objective features of morality that the quasi-realist must tackle in order to save the moral appearances is the possibility of making moral errors. This idea is deeply embedded in moral practise. We think that others are in error and criticize them for making false moral judgments; we think of judgments which seemed right to us at a certain time, that now appears to be mistakes; we wonder whether our present moral attitudes are correct or not. We also think that it is possible to improve our moral attitudes; to make better, not just different moral judgments; we think in terms of moral progress or deterioration. Moreover, if it was not possible to make moral mistakes, both the idea of moral truth and of genuine moral disagreement would be deeply problematic.³³ So, given his ambitious project of accommodating the phenomenology of moral life, Blackburn is committed to give an explanation of moral error. This is not a problem for the realist. She can simply explain moral error as a failing correspondence between moral judgments and independently existing moral facts. For an expressivist however, this way of accounting for moral error is barred.

In the next chapter I will give a discussion of how the quasi-realist can account for first person moral error, the possibility that *I* might be in error. In this chapter I will consider the problems the quasi-realist is facing when ascribing error to others. The worry is that any expressivist view, and therefore quasi-realism as well, is committed to some undesirable form of mind-dependence about ethics.

Blackburn's account of what we do when we accuse others of moral wrongdoing is essentially the same as Ayer's (1946). When we claim that stealing is morally wrong we express higher order attitudes of disapproval towards stealing. We do not try to describe a realm of ethical facts that makes stealing

³³ If Smith says that killing innocent people is morally wrong and Jones says that killing innocent people is morally permissible, and neither of them is wrong, this seems to entail a form of relativism. I discuss this more thoroughly in the next chapter.

wrong. One worry with this view is that it seems to lay a constraint on our criticism of others. When we accuse others of making moral errors, we are merely voicing our own negative emotion towards say stealing and nothing more can be said. Moral objectivity must then be abandoned. In the critical literature about expressivism this worry has been fleshed out in different ways. One is to claim that expressivism leads to relativism. Another worry is that expressivism entails that values are mind-dependent in some objectionable way.³⁴ Both worries share the feature that we are not really justified when talking about moral errors because our moral judgments are in some way dependent upon our attitudes. This in turn would force us to abandon important parts of our moral practise. But that a philosophical theory should make us abandon important parts of our moral practice is implausible. So we should rather abandon expressivism. Early emotivists such as A. J. Ayer and Bertrand Russell admitted that expressivism implies mind-dependence. Russell claimed that “[w]hen we assert that this or that has “value” we are giving expressions to our personal feelings, not to a fact that would be true if our feeling were different” (Russell 1935: 231-2). This is a highly revisionist brand of metaethics. Russell himself accurately captures the uneasiness we feel when confronted with the claim of moral mind-dependence: “I find myself incapable of believing that all that is wrong with wanton cruelty is that I don’t like it.” (Russell, quoted from Wiggins 1987: 185)

It has been argued that Blackburn inherits the problem of mind-dependence from his expressivist predecessors.³⁵ The objectionable form of mind-dependence that quasi-realism is accused of is commonly formulated in conditionals like:

- (1) If I think X is bad, then X is bad.
- (2) If I think X is right, then X is right.

³⁴ Response-dependent theories, for instance, treat morality as mind-dependent. According to theories of this kind, moral goodness (or value or duty) is a disposition to provoke approval in certain people under certain conditions. As such, moral truths are dependent on what people approve of. But response-dependent theories can be fleshed out in very different ways. Both the relevant people and the relevant conditions could be defined with a higher or lower degree of idealization. With enough idealization built into it, a response dependent theory does not have to give a revisionary account of morality. For some different versions of response-dependent theories about values, see Lewis (1989), Johnson (1989) and Smith (1994a).

³⁵ See for example Wiggins (1987) and Kerstein (2006) for claims that quasi-realism implies mind-dependence.

These conditionals are counter-intuitive. We do not normally think of the rightness or wrongness of actions as the products of our thoughts or feelings. The fact that it is wrong to torture babies does not depend upon whatever we think about torturing babies. Thus, if mind-dependence is a consequence of quasi-realism, the quasi-realist project fails by its own standards, namely by failing to save the ethical appearances. What may seem as a dismal disadvantage for the emotivist would be a disaster for the quasi-realist. Accordingly Blackburn denies conditionals like (1) and (2): “We discover such facts, we do not invent them. It is not because of the way we form sentiments that kicking dogs is wrong. It would be wrong whatever we thought about it.” (Blackburn 1984: 217)

The question is whether Blackburn is entitled to this claim. Blackburn often accounts for the objective pretensions of morality by invoking the metaphor of projection.

“[W]e *project* an attitude or habit or other commitment which is not descriptive onto the world, when we speak and think as though there were a property of things which our sayings describe, which we can reason about, know about, be wrong about, and so on. Projecting is what Hume referred to when he talks of “gilding and staining all natural objects with the colours borrowed from internal sentiment”, or of the mind “spreading itself on the world”.” (Blackburn 1984: 171)

The mechanism of projection may explain why moral facts and properties *appear* to be mind-independent, but it does not by itself justify our ordinary understanding of moral mind-independence. On the contrary: projectivism seems to imply mind-dependence. When we are moralizing we are “reacting to a reality which contains nothing in the way of values, duties, rights and so forth” (Blackburn 1981: 164-65). On the projectivist account, moral features are the “children of our sentiments,” not their parents. So, according to the quasi-realist, moral judgments are expressions of attitudes, not representations of independently existing moral facts. Hence, if we had different emotions we would make different moral judgments. To use of one of Blackburn’s recent examples: We do think that parents have an obligation to care for their children. But if our attitudes were different, we might not feel this obligation. From this we can infer three conditionals:

(3) If we did not have these sentiments, we would not be condemning parents who failed to care for their children

(4) If parents X do not have these sentiments, they are likely to fail in caring for their children

(5) If parents X do not have these sentiments, they are under no obligation to care for their children.³⁶

(3) and (4) are uncontroversial, whether one is a moral realist or not. This is because they are not evaluative claims at all. They merely state a causal connection between our sentiments and our actions. The evaluative claim (5), however, is different. It is, according to Blackburn, “false and absurdly so” (Blackburn 2006: 146). But how does Blackburn account for the falsity of (5)? Responding to a situation where we do not care for our children, a realist may simply claim that we are in error: If we had attitudes which allowed for neglecting our children, there would be a mismatch between our attitudes and the relevant moral facts. But this way of saving the mind- independence - by separating psychological states and moral facts - seems to be closed for the quasi-realist; he cannot rely on independently existing moral facts. Still, the quasi-realist would want to maintain that (5) is the expression of a false moral claim. But all he seems entitled to say is that we use the notion of moral rightness or badness to express approval or disapproval, and these judgments are dependent upon our attitudes.

Blackburn objects to this understanding of quasi-realism. He explains the tendency to think that quasi-realism implies mind-dependence as the result of two conflations. First of all it conflates the explanatory and justificatory projects of quasi-realism. In one sense moral judgments is dependent on attitudes: in the sense that moral judgment are explained as expressions of attitudes. But this does not mean that this explanation ought to be any part of the justificatory story anyone might give in support of their moral views. This would be to conflate quasi-realism with subjectivism. According to subjectivism, to make a moral judgment is to *report* our actual attitudes; to say that lying is wrong is just to say that I have a negative attitude towards lying. This makes moral truth straightforwardly mind-dependent, because the moral judgment refers to the

³⁶ These conditionals are slightly rewritten versions of the conditionals mentioned in Blackburn (2006b: 146).

attitudes of the speaker. But this does not follow for the quasi-realist. Although quasi-realism explains moral judgments as expressions of attitudes, this does not limit the *content* of these judgments. Quasi-realism is not reductive, moral judgments need not (and should not, Blackburn would add) be *about* our attitudes.

5.2 The internal reading technique

In *Spreading the Word*, Blackburn gives the following account of statements of mind-dependence:

“Suppose someone said “If we had different sentiments, it would be right to kick dogs”, what could he be up to? Apparently he endorses a certain sensibility: one which lets information about what people feel dictate its attitude to kicking dogs. But nice people do not endorse such a sensibility. What makes it wrong to kick dogs is the cruelty or pain to the animal. That input should yield disapproval and indignation as the output.” (Blackburn 1984: 218)

Blackburn interprets the conditional: “If we had different sentiments, it would be right to kick dogs”, not as a metaethical claim but as an endorsement of a certain sensibility. A moral sensibility, according to Blackburn “is defined by a function from *input* of belief to *output* of attitude.” (Blackburn 1984: 192) If I believe that a dog is being kicked, I can react to this belief by forming a range of different attitudes such as repulsion, or outrage, indifference or amusement. Our sets of moral responses to features of the world make up our moral sensibilities. As discussed in the last chapter, we do not only form attitudes by responding to features of the world, we also respond to the moral sensibilities of others. We find that some moral sensibilities are better than others, some are worthy of endorsement, some should be rejected. By understanding the conditional “If we had different sentiments, it would be right to kick dogs,” as an endorsement of an attitude, Blackburn treats it as a *substantial moral view*. But if the claim to mind-dependence is merely an endorsement of a moral sensibility, the quasi-realist is free to express his repugnance towards this sensibility. This is a crucial step for the quasi-realist. She gives an *internal reading* of the statement of mind-dependence: mind-dependence is not to be understood as a metaethical claim; it is only meaningful as a first-order moral question. The internal reading is a central quasi-realist tactic. According to Blackburn the question of mind-dependence

appears to be a question concerning the metaphysics of morality. But in fact it is not a meta-question at all; it is merely a question of which attitude to adopt towards a certain sensibility.

But of course, this is not the way the accusation of mind-dependence is intended. It is intended not as the expression of a specific moral attitude, but as a theoretical consequence of the explanatory story which the quasi-realist is offering. Blackburn argues that statements of mind-dependence can *only* be understood as substantial moral claims. As a moral view, or as an endorsement of a moral sensibility, the thesis of mind-dependence has no privileged status; from the moral point of view it is merely a repugnant sensibility, which the quasi-realist is free to criticize.³⁷

So in order to meet the accusation of mind-dependence the quasi-realist insists on distinguishing between two stances. As a theorist, the quasi-realist denies that there are any moral facts. From the theorist's point of view, moral judgments are non-cognitive responses to a non-moral world. Our moral judgments do not track moral facts. Involved in the practice of moralizing, however, the quasi-realist is entitled to claim that there are mind-independent moral facts. That it is wrong to torture children does not depend upon whatever we feel about torturing children. It is wrong because of the suffering of the children.

The debate between the quasi-realist and her critics is thus about where to localize claims about mind-dependence. If they are to be understood as metaethical claims, quasi-realism will have to accept mind-dependence. On the other hand, if claims about mind-dependence are to be understood as substantial moral claims, nothing compels the quasi-realist to accept it. If this is right, Blackburn can have his cake and eat it to. On the one hand he will have all the explanatory advantages of an anti-realist. He does not need to enter into the metaphysical quandaries of the realist because the explanation of moral judgments depends on a naturalistically respectable story about emotional reactions. On the other hand, he does not have to struggle with counterintuitive counterfactual statements such as "If we had different sentiments, it would be alright to kick

³⁷ Allan Gibbard (1990: 166-167) gives essentially the same argument for the compatibility of his norm-expressivism and claims of mind-independence.

dogs” because counterfactuals like this are interpreted as substantial moral views, rather than metaphysical consequences of the explanatory story.

Let us for a moment grant Blackburn the internal reading of claims of mind-dependence. We can now see that he is free to apply this technique in a similar manner when discussing relativism.

5.3 The Internal Reading Technique and Relativism

Moral relativism is often understood as the view that moral judgments are true or false relative to some standard which vary between agents and/or groups.³⁸ According to a simplistic version of subjectivism, for instance, the meaning of “X is good” is properly analyzed as “X is approved of by the speaker”. On this analysis, “X is good” and “X is not good” may both be true, depending upon who is making the statement. Construed this way, it is not clear why quasi-realism could be accused of relativism. According to the subjectivist “X is good” is given both a reductive and a descriptive analysis. First, the meaning of moral terms is reduced to mental states. Secondly, the meaning of moral terms is given a descriptive content; when someone says “X is good” he reports her approval of X. But quasi-realism does not attempt to give a reductive analysis of moral judgments. Moreover, for the quasi-realist, moral judgments are understood as expressions of attitudes, not descriptions of them. So if relativism is understood as a consequence of the meaning of moral terms, it does not apply to quasi-realism.³⁹

But if we understand relativism in a somewhat more inclusive manner we can see why accusations of relativism can be directed against the quasi-realist. The problem is that as a theorist, the quasi-realist does not provide us with any objective standards for the justification of our moral judgments. To do so would be to embrace moral realism. Of course, involved in the practice of moralizing she can make normative claims about the standards of justification. But this will just be an expression of her attitudes. There is no reason why everybody else must

³⁸ Although this view is commonly referred to as moral relativism, a more proper name would perhaps be “moral contextualism.” The idea is that there is a hidden indexical variable built into moral statements. “Capital punishment is wrong,” can thus be taken to mean capital punishment is wrong by the standards shared by me and my intended audience. Understood this way moral statements are context sensitive in the same way as indexicals like “I” and “now.” See Brogaard (forthcoming)

³⁹ This is Stevenson’s (1963) way of defending his version of expressivism against relativism. For an exposition of Stevenson’s view and an expressivist defence against charges of relativism, see Horgan and Timmons (2006).

accept these standards, because there is no independent criterion of right and wrong. If you and I are in conflict about some moral issue, all that can be said is that we have conflicting attitudes. From my point of view I will express my disapproval of your attitudes and demand that you change them. But of course, you can do the same towards me. As there are no moral facts, there is no way of settling the issue, without appealing to either your moral sensibilities or to mine.⁴⁰

Blackburn's way of dealing with this problem is to question the standpoint from which the accusation of relativism is made. Someone who claims that quasi-realism implies relativism is insisting that from an external standpoint, outside of our particular moral outlook, there are no criteria for settling a moral conflict. What the relativist demands, according to Blackburn, is for him to give a justification for ethics from a position outside of ethics. "But it is only by using our sensibilities that we judge value. So it is as if we are asked to judge colors with a blindfold on." (Blackburn 1996: 89)

What the quasi-realist needs is a way of saying that some moral judgments are better than others. But this, Blackburn claims, is easily done. Take the claim that women should be educated. Blackburn (1999) imagines a member of the Taliban who holds the reverse. Now a relativist may claim that it is true for the Taliban that women should not be educated, just as it is true for Blackburn that they should. But saying that a judgment is true is just a way of putting the claim in a favourable light.⁴¹ But for Blackburn there is no way that the claim that women should not be educated could be put in a favourable light. The relativist will argue that in that case we have an irresolvable conflict. Blackburn's attitude and the attitude of Taliban are in clash. Neither of the two is able to show that the other is wrong. Blackburn's response is that it is perfectly easy to show that the Taliban is wrong; all he needs to do is to point out the improved lives of educated women. Admittedly, this will perhaps not convince the Taliban. But that does only show that we have no guarantee of convincing people that they are wrong. It does not mean that all moral judgments are equally good.

⁴⁰ Two points are worth noting at this point. First, there might be no ways of settling the matter even if there were moral facts. This will depend on our epistemic access to these facts. Secondly, relativism also seems to threaten the possibility of genuine disagreement. If none of us can be mistaken it will be unclear what we disagree about. Thanks to Knut Olav Skarsaune and Torfinn Huvenes for reminding me of this.

⁴¹ At this point, Blackburn is again arguing from the minimalist theory of truth. Cf. Footnote 32.

Of course, what Blackburn is doing here is to argue from his point of view, using the vocabulary of his ethical outlook. But according to him there is no other way of arguing. To appeal to a standpoint outside our ethical outlook would be to play “the fake externalist game of trying to certify our values without using values” (Blackburn 1996: 89). In effect, Blackburn is applying the same technique as he used to argue against mind-dependence. He argues that relativism should be answered from the perspective of first order moralizing. Just as was the case with the mind-dependence, Blackburn argues that the relativist is making a repugnant *substantial* moral claim; that all moral views are equally good. From within the practice of moralizing Blackburn sees no reason to accept this claim. When the relativist answers that her claim was *about* ethics, and not intended as a substantial moral judgment, Blackburn replies that this claim can only be made from a point of view where all moral attitudes are suspended. As with the question of mind-dependence, relativism, according to Blackburn is, only comprehensible as a statement within ethics. So Blackburn’s replies to the accusations of mind-dependence and relativism rely on essentially the same technique; to give an internal reading of claims that appear to be external. According to Blackburn there is nothing metaphysically demanding in holding that moral judgments are mind-independent and non-relative. It is merely a matter of decent first-order moralizing. Before I discuss the merits of the internal reading technique, I will briefly discuss another argument that Blackburn gives against relativism.

Blackburn argues that the internal reading is enough to fend off relativism but he adds that his claim can be strengthened by the following line of argument:

“We are social animals, with certain biological needs. We have to coordinate our efforts; we have to establish systems of property and promise-keeping and sometimes even government. We can then take comfort in reflecting that there are only so many admirable, coherent, mature, livable ethical systems on offer.” (Blackburn 1998: 308)⁴²

It is worth noticing that this latter strategy of appealing to basic human needs in reality does not strengthen Blackburn’s account at all. Blackburn’s list of the various virtues of ethical systems applies both a normative and non-normative vocabulary. If we focus on the normative laden words such as “mature” and

⁴² For a similar argument, see Blackburn (1984: 197).

“admirable”, it is obvious that these appraisals have to be made on the basis of our ethical outlook. Even given a fixed set of human needs, what is to be counted as mature or admirable is a matter of dispute. So this part of Blackburn’s claim amounts only to a restatement of the internal reading technique; we have to use our moral sensibilities when making moral judgments. To focus on the presumable morally neutral virtues of an ethical system, such as coherence, and liveability, this simply is not enough to ward off the threat of relativism. The potential reward of appealing to coherence and liveability is that these features can be considered to be stance independent; they can be understood in morally neutral ways, so we do not have to rely on a specific outlook to find out whether an ethical system is coherent or livable.

But there are two problems with this line of thought. First of all, although coherence as such is stance-independent, the question of valuing coherence is not. As Blackburn (2002: 176-177) himself argues in a different context, the value of coherence is itself a substantial moral issue. Moreover, coherence is not strong enough to serve as an independent standard for moral justification. The problem with the ethical system of a Nazi- or Stalinist official is not that the ethical systems they approve of are incoherent, but rather (among other things) that they allow for or endorse the suffering of others. Their ethical systems may be inhuman or straightforwardly evil, but not necessarily inconsistent. The same two points can be made about liveability.⁴³ First, the question of whether to value liveability is itself a substantial moral question. There are ethical systems which are simply too stern, or too demanding to support human flourishing.⁴⁴ But again, whether a system of attitudes should support human flourishing is a substantial moral question and not a stance independent fact. Secondly, as was the case with coherence, liveability is not strong enough to provide a moral standard. There is no difficulty in pointing out people who endorsed despicable ethical systems, (and who also committed despicable actions) who nevertheless found their lives very livable. Of course, we can argue that the lives of an Eichman or a Mladic are not livable at all, but this would be a moral assessment based on their

⁴³ Just what Blackburn means by “livable” is not clear to me, but I take it to involve some kind of human flourishing or well being.

⁴⁴ Bjørnstjerne Bjørnson held this view on Christianity, hence the title of his play *Over Ævne*. Susan Wolf’s (1982) criticism of Kantianism and Utilitarianism is based on the claim that these ethical systems are too demanding. I am indebted to Torfinn Huvenes for the reference to Wolf.

evil deeds, not a descriptive, psychological statement. So the quasi-realist is forced to rely on her initial strategy of dealing with relativism and mind-dependence, namely the internal reading technique.

5.4 The Internal Reading and Generality

The internal reading technique, as we have seen, consists in reinterpreting seemingly metaethical claims to substantial moral judgments. But intuitively there is a difference between claims about mind-dependence and substantial moral views. If I say:

(6) It is not the case that if I think killing innocent people is right, then it is right.

This claim seems to have a different status from the claim that

(7) It is wrong to kill innocent people

A natural way of criticising Blackburn's account of mind-independence is therefore to argue that claims like (6) have a different status than claims like (7). (6) seems to express something more than an ordinary moral judgment. One way of fleshing out the difference between claims of mind-dependence and substantial moral views is to focus on the generality of the principle of mind-independence. The quasi-realist understands claims like (6) as substantial moral claims. But mind-dependence does not only concern killing, or kicking dogs, or caring for our children, but potentially every moral judgment. So we need a general formulation of the principle. Suppose Jones utters

(8) The way we think about morality does not determine the moral features of things

How would the quasi-realist understand this claim? According to Blackburn, Jones is referring to a list of substantial moral claims.

According to me “moral truths are mind-dependent” can *only* summarize a list like “if there were no people (or people with different attitudes) then X...” where the dots are filled in by some moral claim about X. One can only assess things on this list by contemplating the nearest possible world in which there are no people or people with different attitudes but X occurs. And then one gives a moral verdict on the situation.” (Blackburn 1998: 311)⁴⁵

This way of accounting for claims about mind-independence is unsatisfactory because it does not provide us with an account of mind-independent moral truths *as such*. Sharon Street (forthcoming) poses the following problem for the quasi-realist. According to Blackburn, Jones’ claim that (8) would summarize a list like: “That capital punishment is morally wrong is not determined by the way we think about capital punishment”; “That abortion is morally permissible is not determined by the way we think about abortion” and so on. But now consider Jones’ neighbour, Smith. He totally agrees with Jones about the correctness of (8); there are independent moral truths. But Smith’s list is quite the opposite of Jones.’ Smith thinks it is a mind-independent moral truth that capital punishment is morally permissible, that abortion is morally wrong, and so on. If Blackburn is right that claims about moral mind-independence *only* can summarize a list of substantial moral claims, he must admit that Smith and Jones are not agreeing about (8). But this is implausible. It seems obvious that Smith and Jones are agreeing about something when both claim that (8).

What are they agreeing about? I think the quasi-realist should reply as follows. Smith and Jones are sharing an attitude towards *the justification* of attitudes; they *disapprove* of justifying moral attitudes by referring to attitudes, such as “Kicking dogs is wrong because I think it is wrong”, and they *approve* of justifying attitudes by referring to what causes the attitudes, for instance “Kicking dogs is wrong because it hurts the dog”. This way of understanding mind-independence is perfectly general because it does not depend on any specific moral judgment. Smith and Jones can disagree on any substantial moral claim but still share the same attitude towards the justification of attitudes.

Construed this way, claims about mind-dependence would be on the same level as other ways of justifying moral judgments. Confronted with someone who

⁴⁵ Blackburn is speaking about mind-dependence and not mind-independence, but this is not significant in this context. As we have seen, expressions of mind-independence are, just as expressions of mind-dependence, taken to be substantial moral views.

argues that moral facts are mind-dependent, the quasi-realist will reply in the same manner as when confronted with, say a moral egoist who justifies all her moral judgments with a reference to her own pleasure or satisfaction. The quasi-realist can, (and will, if she is a nice quasi-realist,) express her contempt towards this way of justifying moral judgments. So the quasi-realist can capture the generality of mind-independence while maintaining that it is a moral claim that is to be understood internal to the practice of moralizing.

I conclude that the quasi-realist is not committed to mind-dependence. In section 5.1 we saw that although the quasi-realist *explains* moral judgments as the expression of attitudes, this does not mean that our attitudes should play a role in the *justification* of our moral judgments. Moreover, the quasi-realist can criticize claims of mind-dependence. This is because the quasi realist interprets mind-dependence as a substantial moral claim, which the expressivist is free to express her repugnance. To the objection that the internal reading implies that mind-independence cannot be given a general formulation, I have argued that Blackburn should understand mind-independence as an attitude towards the justification of moral attitudes. Interpreted this way, the quasi-realist can maintain her claim that mind-independence is to be understood internally to the practice of moralizing, and not as a metaethical question.

6. Quasi Realism and Moral Fallibility

6.1 The Errors of Others and the Errors of my own

6.2 Passing the Buck

6.3 Blackburn on Moral Fallibility

6.4 Egan's Argument

6.5 Biting the Bullet

6.1 The Errors of Others and the Errors of my own

According to the quasi-realist, when we accuse others of moral error, we express our disapproving attitudes. These are formed on the basis of other normative commitments and standards of attitudes that we hold. We must argue from within a certain point of view, there are no external standpoints, according to the quasi-realist. The quasi-realist can give an account of what we do when we accuse others of being guilty of moral error. When we accuse others of error we do this on the basis of what we hold to be morally correct. Similarly, the quasi-realist can account for how I can hold that I made some moral error in the past. I can express my disapproval of the moral judgment that I held before. In that case we just express our disapproval of what we judged earlier from the standpoint that we have at the present moment. The real problem arises when the quasi-realist wants to say that we might be mistaken in our *current* moral views. How does an expressivist make sense of the idea that present moral attitudes may be mistaken? Blackburn acknowledges that this question is harder for the quasi-realist to answer.⁴⁶

“Of course there is no problem thinking that other people may be mistaken, or are indeed mistaken. Anyone thinking that kicking babies for fun is OK is mistaken. The real problem comes with thinking of myself (or of us or our tradition) that I may be mistaken. How can I make sense of fears of my own fallibility?” (Blackburn 1998: 318)

Of course, we are often uncertain about moral issues and say things like: “I don’t know whether capital punishment should be allowed.” This is not problematic for the quasi-realist. In these cases we are just expressing our mixed

⁴⁶ It is also a question that has followed Blackburn throughout his writings. In his first essay on the subject of quasi-realism Blackburn claims that the expression of moral fallibility is “in my opinion the hardest context of all for an anti-realist to deal with” (Blackburn 1993: 127).

attitudes towards capital punishment. But what about cases where we are sure about our attitudes; could Blackburn make sense of the idea that we might be in error? Can the quasi-realist make sense of the fallibility expressed in sentences such as: “I believe that capital punishment is wrong, but I might be mistaken.”? The difficulty is that we do not have a position from which we may disapprove of the attitudes in question, as was the case with accusing others, or our past self of committing a moral error. In this case I believe that capital punishment is wrong, so I do not express any disapproval of my judgment. Still, it is possible that my judgment is wrong. The question the quasi-realist must answer to make sense of this kind of fallibility is this: In virtue of what can our present attitudes be in error? In these cases moral error cannot be accounted for as the expression of a disproving attitude. First of all, because there is nothing in specific to which we might disapprove of. Secondly, because we do not have a point of view from which disapproval might be expressed. So Blackburn needs to give another account of first person present error than he gave of second person error. Before I turn to Blackburn’s preferred solution to this problem, I will consider another way an expressivist may answer this question. The shortcomings of that account will make it clearer what Blackburn needs in order to answer the question of moral error.

6.2 Passing the buck

One possibility open for the expressivist is to “pass the buck” (Timmons 1999: 80). When we are making moral judgments, we form our judgments based on beliefs about the factual world. These beliefs may be mistaken, so the moral error might stem from an error in our non-moral judgments. This is the way C. L. Stevenson (1966) tries to explain moral error.⁴⁷ If we think, as Descartes did, that animals do not have the ability to feel pain, then we might find it morally permissible to kick dogs for fun. This is a moral error. But the error stems from a false belief about dog’s sensitivity to pain. A lot of our moral disagreements do depend upon non-moral beliefs, typically involving the consequences of the some action or event. Many moral issues involve complex factual issues. Two persons

⁴⁷ Here I am following the exposition of Stevenson given in Timmons (1999: 80-81). This is also Blackburn’s solution to the problem in his early essay *Moral Realism* (Blackburn 1993: 111-129, see pages 127-128). Later on however, he develops a subtler account. This account is discussed in the next section.

might agree on a moral standard, say that the right course of action is always to maximize happiness, but still fiercely disagree about whether some particular action maximizes happiness or not. These questions might get very complicated, as in the case of military intervention, health care, or aid to developing countries. But these are factual not normative questions, and as such they pose no problem for the expressivist. She can say that a present moral attitude is in error if it is based on a false factual belief.

False beliefs are not the only source of error which the expressivist can account for. Moral judgment often depends on non-moral reasoning. As Timmons points out (1999: 80) we often draw on deductive and inductive reasoning when thinking about and discussing moral matters. These are other possible sources of error, which the quasi-realist can incorporate. She can say that a present moral attitude is in error, even if it is based on a true factual belief, if the inferential process from those beliefs to the attitude involved some logical mistake.

Now, the expressivist has provided a relatively large room for error on a strictly anti-realist basis, and she can argue that these errors in non-moral judgments are enough to accommodate our commonsensical understanding of moral error. This implies, however, that given that our moral judgments do not depend upon some false beliefs, or erroneous reasoning, it would not be possible to make moral errors. But this lays counterintuitive limitations on what we may criticize in our moral practise. Although errors based on false factual beliefs and errors based on mistaken reasoning might be the things we wonder about when we say things like “Stealing is wrong but I might be mistaken”, we do not limit our criticism of others to point out the errors in their factual beliefs and non-moral reasoning. We also criticise their principles and their attitudes. But if we allow for the possibility that other people might be wrong in their moral principles whereas we ourselves can only be wrong in our non-moral beliefs and reasoning, we have an awkward situation. The solution advocated by Stevenson is to claim that for all people, errors in non-moral facts and reasoning is the only source of error.

There are two main reasons why this account of moral error is unsatisfactory. The first problem is related to moral disagreement. Consider two persons, Smith and Jones, who disagree about capital punishment. The debate about capital punishment is often about factual matters such as the uncertainty of evidence, or the preventive effect of different kinds of punishment. But it need not

be; the matter can be, and often is, discussed as a purely principal issue, independently of questions concerning the possible innocence of the offenders, or the effect on potential law-breakers, or on public opinion. Smith and Jones might agree on all these factual matters and still disagree on the question of capital punishment. Let us also stipulate that their disagreement does not stem from an error in deductive or inductive reasoning. Disagreements of this type—disagreement in principle or disagreement on values – do exist; let us call it a “genuine moral disagreement” (as opposed to “factual disagreement.”) But if Smith says: “Capital Punishment under circumstances C is wrong” and Jones says “Capital Punishment under circumstances C is right” one of them must be in error, even though this error is not a factual error. Two contradictory moral terms cannot both be right. This would imply moral relativism. So Stevenson’s account leads to a dilemma. He can account for moral error in terms of error in factual beliefs. But to do so, he must accept moral relativism. This would be to save one of the objective features of morality – moral error – by rejecting another – moral truth.

The second problem with Stevenson’s answer to the worry about moral error is that it does a poor job in making sense of our commonsensical ideas of moral progress. Moral progress does not seem to consist in merely getting a better grasp of the non-moral facts of a situation. To know more about people’s psychological reactions to an insulting comment, or their physical reaction to a punch in the face does not guarantee any moral progress. We can easily imagine a psychopath learning more about these things without becoming any kinder to the people he meets. He might rather utilize his new knowledge to inflict more harm. Moral progress does not consist in gathering more facts but rather responding to them in a kinder, fairer, or more compassionate way. It is our moral *sensibility* that improves or deteriorates.

6.3 Blackburn on moral fallibility

Genuine moral disagreement is possible, and moral disagreement implies moral error.⁴⁸ The quasi-realist must therefore give another explanation of moral error than the one given by Stevenson. But if moral error is not just a product of factual error or errors in non-moral reasoning, nor a lack of correspondence with

⁴⁸ That is unless we want to accept relativism. In that case it is questionable whether the disagreement is genuine. Cf. footnote 40.

independently existing facts, how does the quasi-realist account for it? Blackburn answers this question as follows:

“If values are projections of a habit of forming some kind of attitude to some kinds of thing, how can I be aware that my *own* attitudes might be defective? [...] The answer is that I know that other people are capable of habits of projection which from my standpoint are deplorable: they judge of things that they are ignorant, and their views are functions of fears and fantasies, blind traditions, prejudice and so on. But then who am I to be sure that I am free of these defects? This thought is quite sufficient to enable me to understand the possibility of my attitudes improving. They ought to be formed from qualities I admire – the proper use of knowledge, real capacity of sympathy, and so on. If they are not, and if the use of those capacities and the avoidance of the inferior determinants of opinion would lead me to change, then the resulting attitude would be not only different but better.” (Blackburn, 1993: 79; see also 1993: 20 and 1998: 318)

Blackburn claims that we can make sense of the idea of moral improvement. Significantly, this possible improvement is not an improvement according to some external standards. The only possible improvement that I would recognize as an improvement, would be to let my attitudes be formed, more than they are presently, on the basis of standards and qualities of which I already approve. Hence, when we are worried that our moral judgment might be mistaken, we do not worry whether they correspond to a set of independently existing moral facts. Rather, we worry about whether our moral judgments would still hold if we measured them against our own standards. Thus, according to Blackburn, *a moral judgment is an error if some improvement of my attitudes, in light of my own standards, would lead me to abandon it.*

This account of moral progress and moral error appeals to an idealization of our attitudes. This provides Blackburn with a room for error: what we presently endorse is not necessarily what is right. To deny this gap would mean to give up the normativity of moral utterances. It is important, however, to note that the idealization which Blackburn makes use of is not merely a question of coherence. If it was, Blackburn’s account of moral error would face the same problems as Stevenson’s. It would not be able to open up for the possibility of genuine moral disagreement. Just as two people who agree on all factual matters might still disagree on some moral issue, so might two people who are absolutely coherent,

fully informed and with maximally unified set of attitudes.⁴⁹ But if they disagree one of them must be in error, unless we want to embrace relativism.⁵⁰

When we wonder whether our attitudes might improve we are not just pondering possible inconsistencies in our system of attitudes. Ideals such as coherence, full information, and maximum unification of our attitudes might be things we admire, but improvement according to these procedural standards are not the only ways of improvement we might undergo. As Blackburn is careful to point out, the standards on which we measure improvement are *substantial* standards, which might themselves be contested. Hence, if I am a romantic who values innocence, I might deplore full information, or if I value engagement in ethical matters I might criticize coherence and unification, because the prize for these virtues might be disengagement (Blackburn 1996: 96; Blackburn 2002: 176-177). In this way Blackburn improves on Stevenson's account by arguing that moral sensibility is the locus for moral improvement. We might improve not only by gathering more information, but also by being more generous, more mature, giving more attention to the needs of others and so on.

There is an obvious problem with this account. What decides whether one of my moral attitudes is erroneous or not is whether it would survive to be measured against the standards I approve of. But this explanation seems just to postpone the problem. Could it not be the case that my own standards, "the qualities I admire," are in error? To answer this question Blackburn has two options. He can either claim that the standards that we presently approve of are immune to error, or that they are not. The first option is implausible. We do change our mind on what should count as an "inferior determinant of opinion", so why should the standards that we presently approve of be counted as immune to error? Fortunately Blackburn opts for the second version: the standards that we presently approve of might also be in error. However, the only way to make sense of this worry is to scrutinize some of our ideals from the standpoint of some of our other ideals. This implies a kind of circularity: our moral beliefs are judged to be

⁴⁹ It might be the case that fully coherent persons with full information and maximum unification would converge in their moral judgements (Smith 1994). But as I argue in the next section, Blackburn's own theory does not have the resources to account for a convergence of moral attitudes, given coherence, full information and maximum unification

⁵⁰ Again, if we embrace relativism, it is a questionable whether the disagreement will be genuine. Cf. footnotes 48 and 40.

erroneous on the basis of ideals which in turn might be revised. But the circularity is not, Blackburn claims, vicious:

“It is true that in saying this I am presupposing one kind of evaluation in giving sense to the possible deficiencies of the other [...]. But this gives nothing an axiomatic status: at the end of a process of reevaluation everything might have changed. The right analogy is that of rebuilding Neurath’s boat” (Blackburn 1993: 79)

This seems to be the only way quasi-realism can explain the notion of moral error without collapsing into realism proper. There is neither a view from nowhere from which we could evaluate our standards, nor a set of independently existing moral facts to which our ideals might correspond. Our set of moral attitudes undergoes a continuous evaluation. This always takes place from within our set of attitudes, but by incorporating a certain idealization of attitudes, it still leaves room for a notion of first person moral error.⁵¹

6.4 Egan’s argument

In a recent article, Andy Egan argues that this account of moral error is flawed. According to Blackburn, I am in moral error if an improving change of attitudes would lead me to abandon a current moral judgment. Now, Egan introduces the concept of a *stable belief*.⁵² “Call a belief *stable* just in case no change that the believer would endorse as an improvement would lead them to change it. Call a belief *unstable* just in case it is not stable.” (Egan 2007: 212) In order for the moral judgment *x* to be a candidate for moral error it must be the expression of an unstable belief; it must be possible for the belief to be revised as a result of an improving change of my attitudes. Next, Egan introduces the concept of *fundamental disagreement*. Smith believes that capital punishment is right; Jones believes that it is wrong. Both Smith’s and Jones’s beliefs are stable;

⁵¹ This seems to be the favoured solution among philosophers who want to save the idea of moral progress and moral error without committing themselves to moral realism. Mark Timmons (1999: 166-170) and Crispin Wright (1992: 200-201) develop very similar accounts.

⁵² It might be confusing that Egan uses the concept of belief when he discusses Blackburn views on moral error. He justifies this by referring to quasi-realists use of minimalism: “Talk of moral belief is brought under the quasi-realist umbrella, as well, in roughly the following way: Believing that stealing is wrong is just taking it to be true that stealing is wrong. What it is to take it to be true that stealing is wrong? Well, it’s to have the attitude that one expresses in sincere utterances of it’s true that stealing is wrong.” Egan (2007: 207) For convenience, I will adopt Egan usage of “belief” when I am discussing his paper.

no improvement would lead them to revise their beliefs. As Egan points out, this might happen in two ways. They can either endorse the same standards of improvement, but start out with different attitudes, or they may start out with similar attitudes but endorse different standards of improvement. Either way, they will end up with different and potentially conflicting stable beliefs. This makes Smith's and Jones' disagreement on capital punishment a *fundamental disagreement*. Capital punishment cannot be both right and wrong, so either Smith or Jones must be guilty of a *fundamental error*: an error in their stable beliefs.

But if the quasi-realist allows for a case like this, a case of fundamental moral error, he is in trouble. For me to be guilty of a fundamental moral error, I must have a belief that is both stable and mistaken. But given Blackburn's account of first person moral error, this is impossible. I can only be in error if my belief is *unstable*, i.e. if an improvement of my attitude would lead me to change my belief. Blackburn's account of moral error thus gives me an a priori guarantee against *fundamental* moral error. Therefore, Egan argues, as a quasi-realist I know *in advance* that there is a certain kind of moral error that others may be guilty of, but of which I am immune.

Egan summarizes his argument in three principles which the quasi-realist is committed to:

1. NO SMUGNESS: There isn't any sort of moral error to which others are subject but against which I have an a priori guarantee of immunity.
2. FUNDAMENTAL FALLIBILITY: It's possible for people's stable beliefs to be mistaken.
3. FIRST PERSON IMMUNITY: I have an a priori guarantee against fundamental moral error. (Egan 2007: 215)

The problem is that one cannot have them all. So, which of the principles should Blackburn deny? Egan claims that neither of the premises can be rejected. To deny the first premise is not an option. That would mean to give up the quasi-realist project of saving the moral appearances. There is no reason to believe that I should have a privileged moral outlook, which makes me immune to the errors that others commit. Blackburn would want to deny the third premise, which is counterintuitive. The problem is that this premise follows from Blackburn's

account of moral error, combined with the second premise, which allows that people's stable beliefs may be mistaken. So, if Blackburn wants to hold on to his account of moral error, he must deny the premise of fundamental fallibility.

What are the prospects of denying this principle? Fundamental fallibility relies on three premises. First, that it is possible to have a stable moral belief. Second, that people may have conflicting stable beliefs. Third, that moral disagreement implies moral error.

In section 6.2 I argued for the last of these premises. Indeed, this is essential for the quasi-realist project of saving our commonsensical notion of moral objectivity. Quasi-realism, as Blackburn is anxious to point out, does not imply relativism. Moreover, quasi-realism promises to save the notion of moral truth. But if so, two contradictory moral views cannot both be true. Consider again the case of Smith and Jones. Smith holds p ; Jones holds not- p . Must the quasi-realist hold that both Smith and Jones are justified in their claims? Blackburn argues that she does not.

“That would require that I myself assent both to p and to not- p , which I don't. If Smith and Jones are each saying that they have found out conflicting things, then at most I allow that one of them is right. I will say that Smith is right if I hold p and that Jones is right if I hold not- p . I shall hold that neither is right if, for instance, I think that we shall not talk in these terms at all.” (Blackburn 1998: 314)

As Blackburn points out, this does not commit him to straightforward moral realism, but is rather an attitude to practical reasoning. (Blackburn 2006: 155) Blackburn's argument is that *from my point of view* I cannot assent to both p and not- p . That would make my system of attitudes incoherent. Blackburn's point is pragmatic. Although two coherent systems of attitude may produce different moral verdicts, I cannot incorporate these conflicting verdicts in my system, as this would make my own system incoherent. From my perspective two conflicting moral judgments cannot meet my approval. Hence, I must claim that either Smith or Jones is in error. So Blackburn is committed to the third premise. Let us therefore consider the first two premises.

Can Blackburn claim that no moral beliefs are stable; that every moral judgment might be revised through some improving change of our attitudes? This

might seem to be a reasonable answer; indeed it seems like just the kind of epistemic modesty that Blackburn is arguing for. But all he can claim is that we *cannot be sure* whether our beliefs would survive an improving change or not. He cannot claim that none of our beliefs would survive an improving change. If this was the case, none of our present moral beliefs would be correct. Recall that Blackburn's view of moral error is that a moral attitude is in error if it would not survive an improving change. To say that none of our present moral beliefs are correct would be just as counterintuitive as to claim that moral error is impossible. So Blackburn is committed to the possibility that some of our moral beliefs are stable.

But perhaps Blackburn can deny the possibility of conflicting stable beliefs? This would mean that if our attitudes were properly idealized, our stable beliefs would converge; fundamental disagreement would be impossible. This is a solution often advocated by realists and constructivists. Michael Smith, for instance, argues that: "The truth of a subject's evaluative beliefs requires that all subjects converge in the desires they would have if they had a set of desires that was maximally informed and coherent and unified" (Smith 2003: 312).⁵³ The first point to make is that such a convergence would not sit well with the expressivist position that Blackburn advocates. One of strengths of this tradition is exactly that it can make sense of fundamental moral disagreement. Moreover, Blackburn's quasi-realism provides us with no reason to suppose that this convergence would happen. Recall that Blackburn's idealization of attitudes is to be understood relative to the standards approved of from a first person perspective. In making sense of first person moral error, Blackburn does not argue for a common standard of idealization, which would apply to every moral agent. As people approve of different standards, their stable moral beliefs will differ as well. As seen in the section on mind-dependence, Blackburn does argue that some standards are better than others and that other people are in error when they do not form their moral attitudes on the basis of standards such as "sensitivity, maturity, imagination and coherence" (Blackburn 1998: 318). But this is how we make sense of *other*

⁵³ Michael Smith's idea of convergence relies on an "Anti-Humean theory of normative reasons". See Smith (1994a: 130- 181). For Blackburn's discussion and rejection of this view, see Blackburn (1998: 263-269).

people's error. In order make sense of the possibility of our own fallibility we have to ask whether our moral attitudes live up to *our own standards*.

The root of the problem posed by Egan lies in a mismatch between our own fallibility and that of others. Blackburn accounts for first person errors relative to the standards that we approve of from the first person perspective, whereas second person errors have no such limitations. We cannot step outside our own evaluations when we ponder our own fallibility and say things like "I might be mistaken". But the fact that someone stably believes that genocide can be justified does of course not stop us from expressing our repugnance at this attitude. We use our own standards when we criticize others, just as we do when we make sense of the idea that we ourselves may be mistaken. But we do not limit our criticism to the error that others would recognize as errors.

6.5 Biting the bullet

The only way out for Blackburn is to bite the bullet. But although Blackburn must accept the argument, he might still argue that Egan has overrated the importance of the problem. Although the quasi-realist account does imply some revisions of our way of thinking about moral error perhaps it does not amount to a full blooded revisionism. If so, the quasi-realist still gives us what we need for thinking and talking about moral error. There are some points that Blackburn might make to justify this stance.

He might argue that the quasi-realist account is more true to the phenomenology of moral fallibility than the account offered by moral realism. According to moral realism it is possible to be wrong in one's stable beliefs. This means that I could make an error that I would never recognize as an error. But this, Blackburn could claim, is not the thing I worry about when I ponder my own fallibility. What I am worrying about is whether some idealized version of me would approve of my actions or not. This might be correct. But a realist might simply answer that if this is the case, the prize for saving the phenomenology is simply too high. If Blackburn wants to hold on to the principle of NO SMUGNESS, he has to deny that it is possible to make errors in stable beliefs. But people might stably believe the most horrible things. So it is highly counter-intuitive to claim that these beliefs cannot be false.

To this Blackburn might point out that the term “stable belief” is very abstract. I do not know which of my beliefs are stable. Hence, I am in an even worse epistemological position to point out stable beliefs, and thus fundamental errors, in other people. There is always the possibility that Smith’s judgment in favour of capital punishment was an unstable, not a stable belief. And Blackburn does not have a problem with errors in unstable beliefs. These are types of errors which others as well as I may commit. Thus Blackburn can argue that I can never find myself in a position to accuse others of a type of error that I myself cannot commit. I would never be in an epistemic situation that allows for this kind of claim. So, for all practical purposes, Blackburn can claim that his notion of moral error holds.

But this detour to the practical sphere does not bring Blackburn very far. Egan’s point is that a quasi-realist can *know in advance* that there is certain kind of error which others may commit but to which I am immune. And this is surely a revisionary consequence of quasi-realism. All that Egan needs for this argument is the theoretical possibility of conflicting stable beliefs, and this, I have argued, is a possibility that Blackburn must allow for.

This is very damaging for the quasi-realist project. The quasi-realist promises to justify our intuitions about moral objectivity. Unless another account of first person moral error can be given, quasi-realism fails by its own standards.

7. Summary and Conclusions

In this thesis I have discussed quasi-realism as an answer to the moral problem. Although it improves on its expressivist predecessor, I have argued that it ultimately fails because it cannot account for the possibility of first person error.

The starting point for the moral problem is two important of moral judgments. First, they aim at objectivity. We normally think there are correct answers to be found on moral matters, and we think that it is possible for our moral judgments to be mistaken. In this respect moral judgments behave as ordinary beliefs. Second, moral judgments are essentially practical. They are action guiding and closely connected to motivation. In this respect they behave as desires. If we combine these two features with a popular view in moral psychology, namely that a belief alone cannot motivate, we have three independently plausible claims, which together appear to be inconsistent. This is what Michael Smith (1994) has called the moral problem. Let me recapitulate the formulation I gave in the introduction:

1. Moral judgments a) aim at objectivity, so b) moral judgments are beliefs.
2. Moral judgments have a necessary connection to motivation.
3. In order to be motivated we need both a relevant desire and a means-end belief. A belief alone is not enough to motivate, and beliefs and desires are distinct existences; so no belief has a necessary connection to motivation.

Early expressivists such as Ayer (1946) and Stevenson (1944) held on to 2 and 3 but denied 1a and 1b. This is a problematic move because it forces us to give up on moral objectivity. Blackburn's quasi-realism intends to improve on its expressivist predecessor by arguing that although moral judgments are expressions of attitudes, the quasi-realist can still justify the objective pretensions of moral discourse. He therefore embraces 1a) but denies 1b), while holding on to 2 and 3.

In order to give an answer to the moral problem, there are two tasks facing the quasi-realist. First, she must show that 2 and 3 give an argument in favour of

quasi-realism. Second, she must show how she can accommodate 1a) while rejecting 1b). The thesis is structured around these two tasks.

The second chapter examined the claim that moral judgments have a necessary connection to motivation. This is motivational internalism. Internalism can be given two different interpretations, a strong and a weak. I argued that the weak version of internalism is the more plausible of the two. However, an expressivist cannot accommodate the weak version. The only way an expressivist can be an internalist is by embracing the strong version of internalism. This is no decisive argument against expressivism; it might be the case that we ought to be strong internalists. Strong internalism is however a far more controversial and counterintuitive version of internalism than weak internalism. So, contrary to what is commonly assumed, expressivism does not have an easy job accounting for moral motivation. But even if we were to accept strong internalism, this will not amount to an argument in favour of expressivism. It might be the case that beliefs can motivate on its own. To reject this possibility the quasi-realist must argue for the Humean theory of motivation (HTM).

This was the subject of the third chapter. I gave a discussion of Michael Smith's teleological argument for HTM. This argument rests on difference in the functional role of desires and beliefs. I argued that on the most charitable reading of this functional account makes it vulnerable to counterexamples. Smith's functional account of beliefs and desires is in need of some refinements. If we assume that this can be given, the most promising way of escaping the Humean claim that belief alone cannot motivate would be to postulate a third kind of mental state, "Besires," in addition to beliefs and desires. The problem with this view is that it, as was the case with expressivism, entails a too close connection between beliefs and motivation. This leaves us in an odd dialectical situation. If we assume that Smith's argument can be properly amended, none of the positions I have discussed can accommodate the weak version of internalism.

The last three chapters dealt with the quasi-realist project of accommodating of our intuitions about moral objectivity. In chapter four I discussed how the quasi-realist can demarcate moral attitudes from non-moral attitudes. There are different possibilities open to the quasi-realist. The moral attitudes might be higher order attitudes, stable attitudes or prescriptive attitudes. Taken one by one none of these features are sufficient. Taken together, however,

they give at least a plausible way of distinguishing moral attitudes. But the plausibility of this suggestion, I argued, relies on the possibility of justifying the attitudes that are expressed in a moral judgment. If not, the moral judgments will be merely manipulative.

This brought us to two potential problems in the quasi-realist account of how moral attitudes can be justified. In the fifth chapter I discussed the accusation that quasi-realism implies that morality is mind-dependent in a way that undermines moral objectivity. Blackburn's reply to this worry is first of all to distinguish between the explanatory and the justificatory stories that the quasi-realist is offering. Although the quasi-realist explains moral judgments as the expression of attitudes, this does not mean that our attitudes should play a role in the justification of our moral judgments. Moreover, the quasi-realist can criticize claims of mind-dependence. This is because the quasi-realist interprets mind-dependence as a substantial moral claim, which the expressivist is free to express her repugnance. To the objection that the internal reading implies that mind-independence cannot be given a general formulation, I argued that Blackburn should understand mind-independence as an attitude towards the justification of moral attitudes. Interpreted this way, the quasi-realist can maintain her claim that mind-independence is to be understood internally to the practice of moralizing, and not as a metaethical question.

The second problem of moral justification that the quasi-realist must answer is this: How is it possible that we ourselves might be mistaken in our moral judgments? Blackburn's answer to this question is that we all have a set of moral standards that we approve of or admire. But we cannot know whether our moral judgments are in accordance with the standards we approve of. When we are worried that our moral judgments might be mistaken, we worry about whether our moral judgments would still hold if we measured them against our own standards. Thus, according to Blackburn, a moral judgment is an error if some improvement of my attitudes, in light of my own standards, would lead me to abandon it. But this account has a serious flaw. As Andy Egan's argument (2007) shows, Blackburn's explanation of first person moral error leads to the unfortunate conclusion that I can know a priori that there is a certain kind of moral error that others may be guilty of, but of which I am immune.

This is not damaging to the *expressivist* project as such. The expressivist is only committed to the claim that moral judgments express non-cognitive attitudes. For the *quasi-realist* project, however, this argument may prove fatal. The quasi-realist is committed to give a non-revisionary account of ordinary moral thought and practise. But as it stands, the quasi-realist explanation of moral error leads to revisionary and counterintuitive results. The possibility of making moral error is an essential part of our commonsensical understanding of morality. If no better account of moral error can be given, the quasi-realist projects fails by its own standards.

In fairness to the quasi-realist it should be noted that the project of articulating a non-revisionary expressivism is still a work in progress, and it is steadily growing more sophisticated. Whether the problem of moral error can be accounted for on a quasi-realist basis is still to be seen.

As an answer to the moral problem, quasi-realism in its present form is unsatisfactory. This, I have argued, is mainly for two reasons. First, it cannot accommodate the most plausible version of internalism. Secondly, until it has provided another account of moral error it does not live up to its project of justifying our intuitions of moral objectivity.

Bibliography:

- Altham, J. E. J. The Legacy of Emotivism. In MacDonald, G and Wright, C. (eds.) *Fact, Science and Morality*. Basil Blackwell, Oxford.
- Anscombe, G. E. M. 1957. *Intention*. Basil Blackwell, Oxford.
- Ayer, A. J. 1946. *Language, Truth and Logic*, Dover Publications, New York
- Blackburn, S. 1984. *Spreading the Word*, Clarendon Press, Oxford
- Blackburn, S. 1993. *Essays in Quasi-Realism*, Oxford University Press, Oxford
- Blackburn, S. 1996. Securing the Nots: Moral Epistemology for the Quasi-Realist in Sinnott-Armstrong and Timmons, M (eds.): *Moral Epistemology*, Oxford University Press, Oxford.
- Blackburn, S: 1998. *Ruling Passions*, Oxford University Press, Oxford.
- Blackburn, S. 1999. Is Objective Moral Justification Possible on a Quasi-realist Foundation? *Inquiry* 42, 213-228.
- Blackburn, S. 2002. Précis of Ruling Passions and Replies. *Philosophy and Phenomenological Research* 65, 122- 135, 164-176
- Blackburn, S. 2006a. Antirealist Expressivism and Quasi-Realism in Copp (ed.): *The Oxford Handbook of Ethical Theory*
- Blackburn, S. 2006b. Must We Weep for Sentimentalism? In Dreier (ed.): *Contemporary Debates in Moral Theory*
- Blair, J. 1995. A Cognitive Development Approach to Morality: Investigating the Psychopath. *Cognition* 57: 1-29.
- Brink, D. 1989. *Moral Realism and the Foundations of Ethics*, Cambridge University Press, Cambridge
- Brogaard, B. Forthcoming. Moral Contextualism and Moral Relativism, *Philosophical Quarterly*, 2008
- Copp, D and Sobel, D. 2001. Against Direction of Fit Accounts of Belief and Desire. *Analysis* 61, 44-53
- Dancy, J. 2000. *Practical Reality*, Oxford University Press, Oxford.
- Divers, J and Miller, A. 1994. Why expressivists bout value should not love minimalism about truth. *Analysis* 54, 12- 19.
- Dreier, J. 1990. Internalism and Speaker Relativism. *Ethics* 101, 6-26
- Egan, A. 2007. Quasi-Realism and Fundamental Error. *Australasian Journal of Philosophy*, Vol. 85, No. 2, 205-219.

- Geach, P. 1960. Ascriptivism. *Philosophical Review* 69, 221-215
- Geach, P. 1965. Assertion. *Philosophical Review* 74, 449-165.
- Gibbard, A. 1990. *Wise Choices, Apt Feelings*, Clarendon Press, Oxford.
- Gibbard, A. 2003. *Thinking How to Live*, Harvard University Press, Cambridge, Massachusetts
- Frankfurt, H. 1971. Free Will and the Concept of a Person. Reprinted in Watson (red.) 2003. *Free Will*, Oxford University Press, Oxford
- Johnson, M. 1989. Dispositional Theories of Value. *Proceedings of the Aristotelian Society* supp vol. 139-174
- Habermas, J. 1991. Von pragmatischen, ethischen und moralischen Gebrauch der praktischen Vernunft. In *Erläuterungen zur Diskursethik*, Suhrkamp. Frankfurt am Main
- Hale, B: 1986. The Compleat Projectivist. *Philosophical Quarterly* 36, 65-84.
- Hale, B: 1993. Can there be a Logic of Attitudes? In Haldane, J and Wright, C: *Reality, Representation, and Projection*, Oxford University Press, Oxford.
- Hare, R: 1961. *The Language of Morals*, Oxford University Press, Oxford.
- Horgan, T and Timmons, M. 2006. Expressivism, Yes! Relativism, No! In Schafer Landau (ed.): *Oxford Studies in Metaethics*, Oxford University Press, Oxford
- Horwich, P. The Essence of Expressivism. *Analysis* 54, 19-20.
- Humberstone, I. L. 1992. Directions of Fit. *Mind*, 101, 59-83
- Hume, D. 1988. *A Treatise of Human Nature*. Oxford University Press, Oxford, 2000.
- Kalderon, M. 2005. *Moral Fictionalism*, Clarendon Press, Oxford
- Kerstein, S. 2006. Reason, Sentiments and Categorical Imperatives. In: Dreier (ed.): *Contemporary Debates in Moral Theory*
- Lewis, D. 1989. Dispositional Theories of Value. *Proceedings of the Aristotelian Society* supp. vol. 113-137
- Mackie, J: 1977. *Ethics. Inventing Right and Wrong*, Penguin Press, London
- McDowell, J. 1978. Are Moral Requirements Hypothetical Imperatives? *Proceedings of the Aristotelian Society*, supp.vol 52, 13-29
- MacIntyre, A. 1981. *After Virtue*, Notre Dame University Press, South Bend, Indiana.
- Miller, A. *An Introduction to Contemporary Metaethics*, Polity Press, Cambridge.
- Nagel, T. 1978. *The Possibility of Altruism*, Princeton University Press, Princeton, New Jersey
- Nichols, S. 2004. *Sentimental Rules*, Oxford University Press, New York

- Platts, M. 1979. *Ways of Meaning*. Routledge, London
- Smith, M. 1994a. *The Moral Problem*, Basil Blackwell, Oxford
- Smith, M. 1994b. Why Expressivists About Value Should Love Minimalism about Truth. *Analysis* 54, 1-11.
- Smith, M. 2001. Some Not Much Discussed Problems for Non-Cognitivism. *Ratio* 14, 93-115.
- Smith, M. 2004. *Ethics and the A Priori*, Cambridge University Press, Cambridge
- Tenenbaum, S. 2006. Directions of Fit and Motivational Cognitivism. In: In Schafer Landau (ed.): *Oxford studies in Metaethics*, Oxford University Press, Oxford
- Russell, B. 1935. *Religion and Science*. Oxford University Press, New York.
- Scanlon, T. 1998. *What We Owe To Each Other*. The Belknap Press of Harvard University Press, Cambridge, MA
- Schafer-Landau, R. 2003: *Moral Realism*, Oxford University Press, Oxford
- Schueler, G. F. 1991. Pro Attitudes and Directions of Fit. *Mind* 100, 277-281
- Svavarsdottir, S. 1999. Moral Cognitivism and Motivation. *The Philosophical Review* 108: 161-219
- Stevenson, C. L. 1937. The Emotive Meaning of Ethical Terms. *Mind*, 46, 65-75
- Stevenson, C. L. 1946. *Ethics and Language*. Yale University Press, New Haven.
- Stevenson, C. L. 1963. *Facts and Values*. Yale University Press, New Haven.
- Stevenson, C. L. 1966. *Ethical Fallibility*. In DeGeorge, R (ed.): *Ethics and Society*
- Stocker, M. 1979. Desiring the Bad: An Essay in Moral Psychology. *The Journal of Philosophy*, 76: 738- 754
- Street, S. Forthcoming. Darwinism, Deep Vindication, and the Schizophrenia of Quasi-Realism about Normativity.
- Taylor, C. 1985. *Human Agency and Language*. Cambridge University Press, Cambridge
- Timmons, M: 1999: *Morality without Foundations*, Oxford University Press, New York
- Tiberius, V. 2000. Humean Heroism: Value Commitments and the Source of Normativity. *Pacific Philosophical Quarterly*, 81, 426-446
- Wallace, J. 1990. How to Argue About Practical Reason. *Mind*, 267-297
- Watson, G. 1975. Free Will. Reprinted in Watson (ed.). 2003. *Free Will*, Oxford University Press, Oxford
- Watson, G. 1987. Responsibility and the Limits of Evil in Schoeman, F. (ed.) *Responsibility, Character and the Emotions*. Cambridge University Press, Cambridge.

- Wiggins, D. 1987. Truth, Invention and the Meaning of Life. In *Needs, Values, Truth*, Clarendon Press, Oxford.
- Wolf, S. 1982. Moral Saints. *Journal of Philosophy*, 79, 419-439.
- Wright, C. 1992. *Truth and Objectivity*, Harvard University Press, Cambridge, MA